

**ANWENDUNG DER GEOMETRISCHEN LÖSUNG DES
GAUSS-MARKOV-PROBLEMS AUF UNVOLLSTÄNDIGE DATEN**

DISSERTATION

zur Erlangung des Grades
eines Doktors der Naturwissenschaften
(Dr. rer. nat.)
der Abteilung für Mathematik
an der Ruhr-Universität Bochum

von

EVA-MARIA MALIN

Bochum 1983

**ANWENDUNG DER GEOMETRISCHEN LÖSUNG DES
GAUSS-MARKOV-PROBLEMS AUF UNVOLLSTÄNDIGE DATEN**

DISSERTATION

zur Erlangung des Grades
eines Doktors der Naturwissenschaften
(Dr. rer. nat.)
der Abteilung für Mathematik
an der Ruhr-Universität Bochum

von

EVA-MARIA MALIN

Bochum 1983

Dissertation eingereicht am: 4.11.1982

Tag der mündlichen Prüfung: 11. 5.1983

Dekan: Prof. Dr. H. Zieschang

Gutachter: Prof. Dr. V. Baumann

Prof. Dr. S. Albeverio

Prof. Dr. A.T. Huckleberry



Inhalt

Seite

Einleitung	5
0. Allgemeine Voraussetzungen, Bezeichnungen	10
1. Gauß-Markov-Problem für das Modell $M(L, Q)$	15
1.0. Existenz und Eindeutigkeit des GM-Schätzers	15
1.1. Berechnung eines GM-Schätzers	18
1.1.0. $Q \in S_+(V)$ regulär	18
1.1.1. $Q \in S_+(V)$ beliebig - "Schrittweise GM-Schätzung"	20
1.1.2. Majorante mit quasiregulärem und regulärem Q	30
1.2. Missing-Data-Problem	35
1.2.0. Formulierung des Problems - Das Modell $M(PL, PQP)$	35
1.2.1. Schätzung von E_{Py} aufgrund der Kenntnis von P_y	38
1.2.2. "Ersetzen fehlender Daten" - der Fall $Q = \sigma^2 I$	40
1.2.3. "Ersetzen fehlender Daten" - $Q \in S_+(V)$ beliebig	44
1.3. Beispiele	55
1.3.0. Regressionsmodell mit festen und zufälligen Effekten: $L = \mathcal{U}(X_1)$, $Q = X_2' X_2 + \sigma^2 I$	55
1.3.1. $\mathcal{U}(X_1) = \mathcal{U}(X_2)$: Fehlende Daten sind "durch Schätzwerte $x_o = AP_y$ ersetzbar"	56
1.3.2. $\mathcal{U}(X_1) \subsetneq \mathcal{U}(X_2)$: Fehlende Daten sind nicht "durch Schätzwerte $x_o = AP_y$ ersetzbar"	62
1.3.3. $\mathcal{U}(X_2) \subsetneq \mathcal{U}(X_1)$: Fehlende Daten sind "durch Schätzwerte $x_o = AP_y$ ersetzbar"	65

	Seite
2. Gauß-Markov-Problem für das Modell $M(L, \Theta)$	68
2.0. Existenz und Eindeutigkeit des GM-Schätzers	68
2.1. Berechnung eines GM-Schätzers	68
2.2. Missing-Data-Problem - Das Modell $M(PL, P\Theta P)$	74
2.3. Beispiele	78
2.3.0. Regressionsmodell mit festen und zufälligen Effekten: $L = \mathcal{M}(X_1)$, $\Theta = \{X_2 \Sigma X_2' + \sigma^2 I \mid \Sigma = \Sigma' \geq 0, \sigma^2 \geq 0\}$	78
2.3.1. $\mathcal{M}(X_1) = \mathcal{M}(X_2)$: GM-Problem lösbar durch $\tilde{G} = PG(I + A)$	79
2.3.2. $\mathcal{M}(X_1) \subsetneq \mathcal{M}(X_2)$: GM-Problem nicht lösbar	79
2.3.3. $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$: GM-Problem lösbar durch $\tilde{G} = PG(I + A)$	80
3. Gauß-Markov-Schätzung in quadratischen Modellen	81
3.0. Zurückführung auf ein lineares Problem	81
3.1. Gauß-Markov-Problem für das Modell $M(L^*, \Theta^*)$	82
3.1.0. Koordinatenfreie Formulierung	82
3.1.1. Das Modell $M(L^*, \Theta^*)$	84
3.1.2. Existenz und Eindeutigkeit des GM-Schätzers	86
3.1.3. Berechnung eines GM-Schätzers	88
3.1.4. Invarianz und Positivität des GM-Schätzers	94
3.2. Missing-Data-Problem - Das Modell $M(P^*L^*, P^*\Theta^*P^*)$	100
3.3. Beispiele	106
3.3.0. Regressionsmodell mit festen und zufälligen Effekten: $L^* = \{X_1 \Sigma_1 X_1' + X_2 \Sigma_2 X_2' \mid \Sigma_1 = \Sigma_1', \Sigma_2 = \Sigma_2'\}$	106
3.3.1. $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$: BQUES für σ_e^2	108
3.3.2. Das Modell " $y_{ij} = a_i + e_{ij}$ "	115
3.3.3. Das Modell " $y_{ij} = \mu + a_i + e_{ij}$ "	122

	Seite
4. Das allgemeine multivariate Regressionsmodell	133
4.0. Das Modell	133
4.1. GM-Schätzer für $E_{\mathbf{p}}\mathbf{y}$ und Σ	133
4.2. Missing-Data-Problem	136
4.3. Beispiel	143
Schlußbemerkung	151
Zeichenerklärung	152
Literaturverzeichnis	154
Lebenslauf	162

Ich danke allen, die mir durch ihr Zuhören, durch Hinweise oder kritische Anmerkungen bei der Erstellung dieser Arbeit geholfen haben.

Insbesondere danke ich Herrn Prof. Baumann für die thematische Anregung sowie für seine kontinuierliche Unterstützung.

Einleitung

Für einen Zufallsvektor y mit Werten in einem endlich dimensionalen reellen linearen Raum V soll unter der Annahme eines allgemeinen linearen Modells $M(L, \Theta)$ - $L := \{E_{Py} \mid P \in \mathcal{P}\}$, $\Theta := \{\text{Cov}_P y \mid P \in \mathcal{P}\} \subseteq S_+(V)$, $\mathcal{P}$ eine vorgegebene Familie von Wahrscheinlichkeitsverteilungen - ein linearer Operator $G: V \rightarrow L$ so bestimmt werden, daß Gy gleichmäßig bester linearer erwartungstreuer Schätzer (BLUE) für E_{Py} $\forall P \in \mathcal{P}$ ist; d.h. in der Klasse aller linearen erwartungstreuen Schätzer soll $\text{Cov}_P Gy$ $\forall Q \in \mathcal{Q}$ gleichmäßig minimiert werden. (Gauß-Markov-Problem)

Grundlage dieser Arbeit ist das sogenannte Gauß-Markov-Theorem, das in seiner allgemeinsten Formulierung (siehe z.B. Lemma 0(3) bzw. DRYGAS 1971, Theorem 1.4.) folgende anschauliche Bedeutung hat:

Zu dem Problem einen Schätzer minimaler Varianz zu bestimmen gibt es ein duales geometrisches Problem, nämlich in V den Punkt $l \in L$ zu ermitteln, der zum Meßwert y den kleinsten Abstand hat, wobei der Abstandsbegriff durch eine durch Θ induzierte Norm bzw. Seminorm gegeben ist.

(Eine genauere Beschreibung dieses Dualitätsprinzips findet sich bei DURBIN & KENDALL 1951.)

Gesucht ist eine Darstellung von V als direkte Summe von Teilräumen $V = L \oplus W$. Gy ist dann als Projektion von y auf L entlang W zu berechnen. (Geometrische Lösung)

In Anknüpfung an Arbeiten von Kruskal (KRUSKAL 1961, 1968), Eaton (EATON 1970, 1978) und Drygas (DRYGAS 1970 ff) gehe ich von einer koordinatenfreien Formulierung aus. Dieses ermöglicht zunächst (Kapitel 1.1.) eine systematische Übersicht über in der Literatur angegebene Darstellungen von GM-Schätzern bei beliebiger (bekannter) Kovarianz Q -

$\Theta = \{\sigma^2 Q\}$ -, insbesondere eine Beantwortung der Frage, ob tatsächlich dieselben Multiplikationsmatrizen (Projektoren) vorliegen oder nur f.s. gleiche Schätzer.

Da durchgehend rein geometrisch, d.h. ohne Berücksichtigung der jeweiligen stochastischen Bedeutung, argumentiert wird (es werden nur Projektoren auf bestimmte lineare Teilräume berechnet), sind die Ergebnisse auf ein stochastisch völlig anderes Problem anwendbar: das Ersetzen fehlender Daten bei unkorrelierten Einzelmessungen - insbesondere für $Q = \sigma^2 I$ - (Kapitel 1.2.)

Kruskal hat bereits eine koordinatenfreie Lösung dieses Problems angegeben (KRUSKAL 1961). Der von mir beschrittene Weg führt unmittelbar auf eine geschlossene Darstellung des zu berechnenden Projektors.

Daß dieses für die empirische Forschung wichtige Verfahren des Ersetzens fehlender Daten nicht bei beliebig korrelierten Beobachtungen angewendet werden kann, ist ein zentrales Ergebnis dieser Arbeit. (Notwendige und hinreichende Bedingungen werden in Theorem 1(8) angegeben.)

Die Zahlenbeispiele (Kapitel 1.3.) verdeutlichen zum einen den numerischen Vorteil dieser Methode gegenüber einer direkten Berechnung (Ersparnis von Rechenzeit und Speicherplatz durch Inversenbildung kleinerer Matrizen) und zum anderen die Grenzen ihrer Anwendbarkeit: Es werden tatsächlich nicht in jedem Fall die (existierenden) optimalen Schätzer berechnet.

I.a. sind die tatsächlichen Korrelationen der Einzelmessungen nicht bekannt. In vielen Situationen erscheint es sinnvoll, nach Schätzern zu suchen, die bezüglich einer ganzen Klasse von Kovarianzoperatoren gleichmäßig optimal sind:

Insbesondere beim Schätzen von Varianzkomponenten (Kapitel 3.) sowie in multivariaten Modellen (Kapitel 4.) aber auch in zahlreichen anderen Situationen bieten sich bestimmte Annahmen über die Struktur der Kovarianzmatrix an.

Um die in Kapitel 1. erzielten Ergebnisse auf eine beliebige Menge $\Theta \subseteq S_+(V)$ anwenden zu können, zeige ich in 2.1., daß die Existenz eines GM-Schätzers für $M(L, \Theta)$ gleichbedeutend damit ist, daß $M(L, \Theta)$ zu einem Modell $M(L, \bar{Q})$ mit $\bar{Q} \in S_+(V)$ "äquivalent" ist.

Das Konzept der Äquivalenz- (und Ordnungs-) Relation zwischen linearen Modellen ermöglicht es zum einen, die von Eaton (EATON 1978, Theorem 3.1.) und Drygas (DRYGAS 1976, Theorem 2.1.) angegebenen Existenzbedingungen als Symmetrieeigenschaft linearer Operatoren zu formulieren, so daß sie numerisch leicht nachprüfbar sind. Zum anderen liefert die Anwendung von Kapitel 1.1. auf $M(L, \bar{Q})$ unmittelbar eine Darstellung des optimalen Schätzers für $M(L, \Theta)$ (falls vorhanden).

Im Falle unvollständiger Daten stellt sich die Frage der Existenz gleichmäßig bester Schätzer von neuem. In Kapitel 2.2. zeige ich, daß man mit Hilfe desselben $\bar{Q}$ wie für das Ausgangsmodell $M(L, \Theta)$ prüfen kann, ob ein optimaler Schätzer existiert und wie man ihn gegebenenfalls berechnet.

Dabei zeigt sich ein enger Zusammenhang mit dem in 1.2. behandelten Problem des Ersetzens fehlender Daten. Letzteres kann unter bestimmten Voraussetzungen als ein spezielles Existenzproblem gleichmäßig bester Schätzer aufgefaßt werden. Ein Vergleich der Zahlenbeispiele 1.3. und 2.3. macht diesen Zusammenhang besonders deutlich.

Zur Berechnung bester quadratischer Schätzer für Varianz- und Kovarianzkomponenten werden in Kapitel 3. unter der Annahme einer Quasinormalverteilung für die Beobachtungen y die Ergebnisse aus Kapitel 2. auf lineare Modelle $M(L^*, \theta^*)$ in $Y := yy'$ angewendet. Die von Seely (SEELY 1971) und Drygas (DRYGAS 1977) angegebenen Existenzbedingungen für gleichmäßig beste quadratische Schätzer ergeben sich unmittelbar als Spezialfälle der Existenzbedingungen für lineare Schätzer.

Eine explizite Darstellung (lokal) bester erwartungstreuer quadratischer Schätzer wird entwickelt und der Zusammenhang mit Raos "MINQUE-Theorie" hergestellt. Als Folge des modellmäßigen Zusammenhangs zwischen $M(L, \Sigma)$ und $M(L^*, C^*)$ (Quasinormalverteilung!) zeigt sich, daß die Invarianz dieser Schätzer bezüglich Translationen von $E_P y$ genau dann gesichert ist, wenn ein gleichmäßig bester linearer Schätzer für $E_P y$ existiert.

Falls einzelne Meßdaten fehlen, ist es vielfach möglich, unter Verwendung der Ergebnisse aus Kapitel 1.2. und 2.2. die neuen Schätzfunktionen mit Hilfe einer erweiterten "Beobachtungsmatrix" $Pyy'P + B$ zu ermitteln (Kapitel 3.2.). Dieses Verfahren erfordert lediglich das Lösen eines linearen Gleichungssystems und ist vermutlich meist schneller als eine direkte Berechnung des neuen Projektors.

Die Beispiele 3.3.2. und 3.3.3. zeigen jedoch, daß man i.a. nur für spezielle Parameterwerte lokal beste Schätzer erhält.

Ein unmittelbares "Ersetzen der fehlenden Meßwerte" (Ergänzen von $P y$) führt nur in Ausnahmefällen zu (lokal) besten (invarianten) quadratischen Schätzern (Beispiele 3.3.1. und 4.3.).

Für das Beispiel des multivariaten Regressionsmodells (Kapitel 4.) stellt dieses Verfahren jedoch eine Verallge-

meinerung der Wilks-Methode zum Ersetzen fehlender Daten dar (WILKS 1932), die für den Fall schwach korrelierter Einzelmessungen aufgrund empirischer Untersuchungen sowie von Simulationsstudien (TIMM 1970, GLEASON & STAELIN 1975) und nicht zuletzt wegen ihrer einfachen Handhabung Eingang in die sozialwissenschaftliche Praxis gefunden hat (MEYERS, Time-Sharing Computation in the Social Sciences 1973, NIE et.al., Statistical Package for the Social Sciences).

Im Rahmen dieser Arbeit läßt sich begründen, daß man - Quasinormalverteilung vorausgesetzt - für den Fall unkorrelierter Beobachtungen ($\Sigma = (\sigma_{ij})$ diagonal) die besten invarianten erwartungstreuen quadratischen Schätzer für σ_{ij} bzw. (für $i=j$) die besten erwartungstreuen quadratischen Schätzer überhaupt erzielt (Beispiel 4.3.).

0. Allgemeine Voraussetzungen, Bezeichnungen

Sei $(\Omega, \mathcal{A})$ ein Meßraum,
 $\mathcal{P}$ eine Familie von Wahrscheinlichkeitsmaßen,
 V ein endlich-dimensionaler Vektorraum über $\mathbb{R}$
mit Skalarprodukt $(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$.

Ferner sei $y: \Omega \rightarrow V$ eine Abbildung, so daß $\forall x \in V$
 (x, y) eine reelle Zufallsvariable ist.

Ich nehme an, daß $\forall x, z \in V, \forall P \in \mathcal{P}$
 $E_P(x, y)$ und $\text{Cov}_P((x, y), (z, y))$ existieren.

Lemma 0(1) i) $\forall P \in \mathcal{P}$ genau ein $\mu = \mu(P) \in V$:
 $E_P(x, y) = (x, \mu) \quad \forall x \in V$.
ii) $\forall P \in \mathcal{P}$ genau ein $Q = Q(P) \in S_+(V)$:
 $\text{Cov}_P((x, y), (z, y)) = (x, Qz) \quad \forall x, z \in V$.
Bezeichnung: $\mu =: E_P y$ und $Q =: \text{Cov}_P y$.

Beweis: mit Riesz'schem Darstellungssatz
für lineare Funktionale
(siehe auch KRUSKAL 1961)

Lineare Hypothese: Es gibt einen linearen Teilraum $L \subseteq V$:
 $E_P y \in L \quad \forall P \in \mathcal{P}$.

Ich nehme an, daß i) $L = \{ \mu = \mu(P) \mid P \in \mathcal{P} \}$
ii) $\{ Q = Q(P) \mid P \in \mathcal{P} \} =: \mathcal{Q} \subseteq S_+(V)$ (*)

Die Menge aller Abbildungen $y: \Omega \rightarrow V$ wie oben, mit $E_P y \in L$
und $\text{Cov}_P y \in \mathcal{Q}$ werde im Folgenden als das lineare Modell
 $M(L, \mathcal{Q})$ bezeichnet. Falls $\mathcal{Q} = \{Q\}$ einelementig: $M(L, Q)$.

Gauß-Markov-Problem

Zum Modell $M(L, \Theta)$ (L, Θ bekannt) wird ein linearer Operator $G: V \rightarrow L$ gesucht, so daß $\forall P \in \mathcal{P}$ Gy gleichmäßig bester linearer erwartungstreuer Schätzer für $E_P y$ ist; d.h.

- i) $E_P(Gy, x) = E_P(y, x) \quad \forall x \in V, P \in \mathcal{P}$
 - ii) $\text{Var}_P((Gy, x)) = (x, GQG'x) \leq (x, GQG'x)$
 $\forall x \in V, P \in \mathcal{P}$ und $\forall G: V \rightarrow L$ die i) erfüllen
- (Wegen $E_P(Gy, x) = (E_P y, G'x)$ und Annahme $(*)$ ist i) äquivalent zu " $G1=1 \quad \forall 1 \in L$ ".)

Gy heißt dann Gauß-Markov-Schätzer (GMS) für $E_P y$ im Modell $M(L, \Theta)$.

G heißt Gauß-Markov-Operator (GMO) zum Modell $M(L, \Theta)$. (vergleiche EATON 1978)

Lemma 0(2) G ist GMO zum Modell $M(L, \Theta)$

$\Leftrightarrow$

(Gy, x) ist gleichmäßig bester linearer erwartungstreuer Schätzer (BLUE) für alle linearen Funktionale $(E_P y, x)$, $x \in V$.

Beweis: siehe DRYGAS 1970, Theorem 3.9.
und Theorem 3.16.

Zum Nachweis, daß ein gegebener Operator $G: V \rightarrow V$ GMO zum Modell $M(L, \Theta)$ ist, eignet sich besonders die Projektions- und die Symmetrie-Eigenschaft von G :

Lemma 0(3) (Projektionskriterium)

$G: V \rightarrow V$ ist GMO zum Modell $M(L, \Theta)$

$\Leftrightarrow$

- i) $G1 = 1 \quad \forall 1 \in L$ und
- ii) $Gx = 0 \quad \forall x \in \sum QL^\perp$

Beweis: folgt unmittelbar aus DRYGAS 1971,
Theorem 1.4.

(vergleiche EATON 1978, Theorem 2.2. und
RAO 1973, Theorem 3.1., Corollar (3.9))

Corollar 1 G ist GMO zum Modell $M(L, \Theta)$
 $\Rightarrow$ G ist Projektor (G ist idempotent).

Bemerkung: G ist i.a. nicht symmetrisch, d.h. G ist i.a.
nicht orthogonale Projektion.

Als notwendige und hinreichende Bedingung für die Existenz
eines GMO ergibt sich unmittelbar:

Corollar 2 $\exists G: V \rightarrow L$, GMO zum Modell $M(L, \Theta)$
 $\Leftrightarrow L \cap \sum QL^\perp = 0$.

Beweis: folgt aus Lemma 0(3), bzw. DRYGAS
1976, Theorem 2.1.

(vergleiche EATON 1978, Theorem 3.1. und
RAO 1971, Theorem 5.5.ii))

Lemma 0(4) (Symmetriekriterium)

$G: V \rightarrow L$ idempotent, so gilt:

G ist GMO zum Modell $M(L, \Theta)$

$\Leftrightarrow GQ$ symmetrisch $\forall Q \in \Theta$.

Beweis: siehe DRYGAS 1977, Corollar 2.2.c)
(vergleiche RAO 1971, Theorem 5.1.)

Corollar: $G: V \rightarrow L$ idempotent, so gilt:

- i) G ist symmetrisch $\Leftrightarrow G = P_L$,
 ii) P_L ist GMO zu $M(L, \Theta) \Leftrightarrow QL \subseteq L \quad \forall Q \in \Theta$

Beweis: i) G symmetrisch

$$\Leftrightarrow (L.0(4)) \quad G \text{ ist GMO zu } M(L, I)$$

Da I regulär, folgt $G = P_L$

ii) Wegen L.0(3): P_L ist GMO zu $M(L, \Theta)$

$$\Leftrightarrow QL^\perp \subseteq L^\perp \quad \forall Q \in \Theta.$$

$$QL^\perp \subseteq L^\perp \Leftrightarrow QL \subseteq L.$$

(vergleiche EATON 1978, Theorem 3.3.,

REHDER 1980, Theorem 3.1.)

Lemma 0(5) Seien $\Theta, \tilde{\Theta} \subseteq S_+(V)$.

Durch $M(L, \Theta) \equiv M(L, \tilde{\Theta}) : \Leftrightarrow$

$$\sum QV = \sum \tilde{Q}V \quad \text{und} \quad \sum QL^\perp = \sum \tilde{Q}L^\perp$$

wird eine Äquivalenzrelation R auf der Menge

$$\mathcal{M}_{L,V} := \{ M(L, \Theta), \Theta \in S_+(V) \}$$

definiert.

Beweis: R ist reflexiv, symmetrisch und transitiv.

Lemma 0(6) Seien $\Theta, \tilde{\Theta} \subseteq S_+(V)$.

Durch $M(L, \Theta) \leq M(L, \tilde{\Theta}) : \Leftrightarrow$

$$\sum QV \subseteq \sum \tilde{Q}V \quad \text{und} \quad \sum QL^\perp \subseteq \sum \tilde{Q}L^\perp$$

wird eine Ordnungsrelation auf der Menge

der Äquivalenzklassen $\mathcal{M}_{L,V}/R$ definiert.

(Dabei stehe $M(L, \Theta)$ für die Klasse $[M(L, \Theta)]$.)

Beweis: Die Relation ist reflexiv, identitiv und transitiv.

Corollar 1 $\theta \in S_+(V)$, so gilt:

$$M(L, \theta) \leq M(L, I) \iff QL \subseteq L \quad \forall Q \in \theta.$$

Beweis: siehe Corollar zu Lemma 0(4).

Corollar 2 $\theta, \tilde{\theta} \in S_+(V)$, $\theta \subseteq \tilde{\theta} \implies M(L, \theta) \leq M(L, \tilde{\theta})$.

Die Umkehrung gilt offensichtlich nicht.

Lemma 0(7) (Majorantenkriterium)

$G: V \rightarrow L$ GMO zum Modell $M(L, \tilde{\theta})$ und

$$M(L, \theta) \leq M(L, \tilde{\theta})$$

$$\implies G \text{ ist GMO zum Modell } M(L, \theta).$$

Beweis: folgt aus Lemma 0(3).

(vergleiche RAO 1971, Theorem 5.2.)

Lemma 0(8) (Farkas' Theorem)

Seien V_1, V_2 lineare Räume, $M \subseteq V_2$ ein linearer Teilraum und $X: V_1 \rightarrow V_2$ eine lineare Abbildung, so gilt:

$$i) \quad (X^{-1}(M))^{\perp} = X'M^{\perp}$$

$$ii) \quad (X'M^{\perp})^{\perp} = X^{-1}(M)$$

Beweis: siehe DRYGAS 1970, (2.28)

Corollar Seien M_1 und M_2 lineare Teilräume von V , so gilt:

$$i) \quad M_1^{\perp} + M_2^{\perp} = (M_1 \cap M_2)^{\perp}$$

$$ii) \quad (M_1 + M_2)^{\perp} = M_1^{\perp} \cap M_2^{\perp}$$

Beweis: siehe DRYGAS 1970, (2.29)

1. Gauß-Markov-Problem für das Modell $M(L, Q)$

1.0. Existenz und Eindeutigkeit des GM-Schätzers

Ich nehme an, daß $\forall P \in \mathcal{P} \text{ Cov}_P Y = \sigma^2 Q$ ($\sigma^2 = \sigma^2(P)$) ist. Dabei sei $Q \in S_+(V)$ ein bekannter linearer Operator.

Für $\Theta := \{\sigma^2 Q \mid \sigma^2 \geq 0\}$ ist $M(L, \Theta) = M(L, Q)$, so daß ein GMO zum Modell $M(L, Q)$ zu bestimmen ist.

Lemma 1.0(1) Für beliebiges $Q \in S_+(V)$ ist $L \cap QL^\perp = 0$.

Beweis: Sei $l = Qa \in L$, $a \in L^\perp$, so ist

$$0 = (l, a) = (Qa, a), \quad 0 = Qa = l, \quad \text{da } Q \in S_+(V).$$

Corollar: Für beliebiges $Q \in S_+(V)$ ist die Existenz eines GMO zum Modell $M(L, Q)$ gesichert.

Beweis: folgt aus Lemma 0(3)

Eindeutigkeit des GM-Schätzers:

Lemma 1.0(2) $Q \in S_+(V)$ regulär $\Rightarrow L \oplus QL^\perp = V$

Beweis: Sei $x \in (L + QL^\perp)^\perp = L^\perp \cap Q^\perp L$, so ist

$$0 = (x, Qx) \quad \text{und} \quad x = 0,$$

$$\text{also } V \subseteq L + QL^\perp \subseteq V.$$

Mit Lemma 1.0(1) folgt die Behauptung.

Bemerkung: $QL^\perp = L^{\perp \circ}$, orthogonales Komplement von L bezügl. des durch $(\cdot, \cdot)_\circ := (\cdot, Q^\perp \cdot)$ auf V definierten Skalarproduktes.

Corollar: Für reguläres $Q \in S_+(V)$ ist der GMS G_y für E_{PY} im Modell $M(L, Q)$ für alle $y \in V$ eindeutig bestimmt.

Bemerkung: $G = P_L^\circ$, orthogonale Projektion auf L
bezüglich $(\cdot, \cdot)_\circ$.
Insbesondere ist G symmetrisch bezügl.
 $(\cdot, \cdot)_\circ$. (vergleiche Lemma 0(4), Corollar).

Lemma 1.0(3) $Q \in S_+(V)$ beliebig $\Leftrightarrow V = QV \oplus K(Q)$.

Beweis: $(QV)^\perp = Q^\perp(V^\perp) = Q^\perp(0) = K(Q)$.
(siehe Lemma 0(8), "Farkas' Theorem")

Lemma 1.0(4) Sei Q^- beliebige g -Inverse zu Q ($QQ^-Q = Q$).
Durch $(\cdot, \cdot)_\circ := (\cdot, Q^- \cdot)_\circ$ wird auf QV ein
Skalarprodukt definiert.
Der Wert des Skalarproduktes ist unabhängig
von der speziellen Wahl des Q^- .

Beweis: Für $x \in QV$ ist $QQ^-x = x$, $Q^-x = 0$ nur für
 $x = 0$. Q^- auf QV eindeutig.

Bemerkung: $Q \in S_+(V)$, so gibt es immer eine symmetrische
 g -Inverse Q^- . Insbesondere ist $Q^+ \in S_+(V)$.
In Abschnitt 1.3.1. wird eine Methode zur
Berechnung einer symmetrischen g -Inversen
angegeben.

Lemma 1.0(5) $QV = (L \cap QV) \oplus QL^\perp$

Beweis: $(L \cap QV) + QL^\perp \subseteq QV$,
 $[(L \cap QV) + QL^\perp]^\perp = (L^\perp + K(Q)) \cap Q^\perp L \subseteq K(Q)$,
denn für $x = u + v \in Q^\perp L$ ($u \in L^\perp$, $v \in K(Q)$)
ist $0 = (Qx, u) = (Qx, x)$, $x \in K(Q)$.
Mit Lemma 1.0(1) folgt die Beh.

Bemerkung: $QL^\perp \perp_\circ (L \cap QV)$ bezügl. $(\cdot, \cdot)_\circ$.

Lemma 1.0(6) $L + QV = L \oplus QL^\perp$.

Beweis: Für $x \in (L + QV)^\perp = L^\perp \cap Q^{-1}L$ ist

$$(Qx, x) = 0, \quad x \in K(Q) \cap L^\perp;$$

$$\text{also } L + QV \subseteq L + QL^\perp \subseteq L + QV.$$

Mit Lemma 1.0(1) folgt die Behauptg.

Corollar 1 Für $y \in L + QV$ ist der GMS Gy im Modell $M(L, Q)$ eindeutig bestimmt.

Beweis: folgt aus Lemma 0(3)
(Projektionskriterium).

Corollar 2 $V = [L \oplus QL^\perp] \oplus (L^\perp \cap K(Q))$

Ich beschränke mich in dieser Arbeit auf die Betrachtung von Projektoren $G: V \rightarrow L$ entlang $W := QL^\perp \oplus (L^\perp \cap K(Q))$.
 G hängt offensichtlich von dem gegebenen Skalarprodukt ab. Sei etwa $V = (L + QV) \oplus M$, $H: V \rightarrow L$ Projektor auf L entlang $QL^\perp \oplus M$, so ist $H \neq G$.

Für einen beliebigen GMO $H: V \rightarrow V$ ist jedoch $(G-H)Q = 0$,
also $\text{Cov}_P (G-H)y = 0$, $(G-H)y = (G-H)E_P y = 0 \quad \forall y \in V, \forall P \in \mathcal{P}$, d.h.
 $Gy = Hy$ P-f.s. $\forall P \in \mathcal{P}$.

1.1. Berechnung eines GM-Schätzers

1.1.0. Q regulär oder quasiregulär bezüglich L

Definition: Q heiÙe quasiregulär bezüglich L
: $\langle \Rightarrow \rangle \quad L \subseteq QV$.

Theorem 1(1) $Q \in S_+(V)$ regulär, $G:V \rightarrow L$ GMO zum Modell $M(L,Q)$,
so gilt:

- i) $G = P_L^\circ$, orthogonale Projektion auf L
bezügl. $(.,Q^{-1}.)$.
 $\forall y \in V$ ist $l_0 = Gy$ die eindeutig bestimmte
Lösung des Extremalproblems

$$(y-l, Q^{-1}(y-l)) \rightarrow \min_{l \in L}$$

- ii) $G = P_L \quad \langle \Rightarrow \rangle \quad QL = L$

Beweis: i) folgt aus Lemma 1.0(2) und
Lemma 0(3)

ii) siehe Corollar zu Lemma 0(4),
bzw. KRUSKAL 1968, Theorem 1.

Theorem 1(2) $Q \in S_+(V)$ quasiregulär bezügl. L, $G:V \rightarrow L$
Projektor auf L entlang $W := QL^\perp \oplus K(Q)$,
so gilt:

- i) $G = P_L^\circ P_{QV}$, P_L° orthogonale Projektion
auf L bezüglich $(.,Q^{-1}.)$ in QV.
 $\forall y \in V$ ist $l_0 = Gy$ die eindeutig bestimmte
Lösung des Extremalproblems

$$(y-l, Q^+(y-l)) \rightarrow \min_{l \in L}$$

$$\text{ii) } G = P_L \quad \langle \Rightarrow \rangle \quad QL \subseteq L.$$

Beweis: i) folgt aus Lemma 1.0(5) und
Lemma 0(3)

ii) siehe Corollar zu Lemma 0(4)

Corollar: (Verallgemeinerte Aitken-Formel)

Sei U ein reeller Vektorraum mit Skalarprodukt,
 $A: U \rightarrow V$ eine lineare Abbildung,
 $L = AU \subseteq QV$, $Q \in S_+(V)$,
 G der Projektor aus Theorem 1(2),
so gilt:

$$G = A(A'Q^+A)^-A'Q^+$$

Beweis: Sei $\bar{G} := A(A'Q^+A)^-A'Q^+$,

$$\bar{G}V \subseteq AU = L.$$

Wegen $A(A'Q^+A)^-A' = A(A'Q^+A)^+A'$ für
eine beliebige g -Inverse $(A'Q^+A)^-$
ist $\bar{G}$ idempotent.

Wegen $L \subseteq QV$, $A' = 0$ auf $L^\perp \supseteq K(Q)$ ist
 $\bar{G}Q = A(A'Q^+A)^-A'Q^+Q = A(A'Q^+A)^-A'$
symmetrisch.

Mit Lemma 0(4) (Symmetriekriterium)
ist $\bar{G}$ GMO zu $M(L, Q)$.

Wegen $\bar{G} = 0$ auf $K(Q)$ ist $\bar{G} \equiv G$.

Anwendung: i) Für $L = \mathcal{N}(X) \subseteq \mathbb{R}^n$, X bekannte reelle Matrix,
ergibt sich die bekannte Aitken-Formel.

ii) Für einen beliebigen Teilraum $L \subseteq V$ ist
 $L = P_L V$. Mit $U = V$, $A = A' = P_L$ ergibt sich

$$G = P_L (P_L Q^+ P_L)^- P_L Q^+ = [P_L Q^+ P_L]^+ Q^+$$

1.1.1. $Q \in S_+(V)$ beliebig - "Schrittweise GM-Schätzung"

Konstruktionsprinzip

Analog zu $QV = (L \cap QV) \oplus QL^\perp$ (Lemma 1.0(5)) kann man L zerlegen in einen "regulären" Anteil $L_2 := L \cap QV$ und einen "singulären" Anteil L_1 , so daß $L = L_1 \oplus L_2$.

Dann ist $V = L_1 \oplus (L \cap QV) \oplus QL^\perp \oplus (L^\perp \cap K(Q))$. $y \in V$ hat dann eine eindeutige Zerlegung in $y = l_1 + l_2 + u + v$ (mit $l_1 \in L_1$, $u \in QL^\perp$ und $v \in L^\perp \cap K(Q)$).

Es gibt mindestens zwei verschiedene Möglichkeiten, $l = Gy$ "schrittweise" zu berechnen:

- i) $l_1 = G_1 y$ ist Projektion von y auf L_1
 entlang $QV \oplus (L^\perp \cap K(Q))$,
 $l_2 = G_2 (y - l_1)$ ist Projektion von $y - l_1$ auf $L_2 = L \cap QV$
 entlang $QL^\perp \oplus (L^\perp \cap K(Q))$.

$$Gy = l = l_1 + l_2 = [G_1 + G_2(I - G_1)] y. \quad (\text{Theorem 1(3)})$$

Dabei lassen sich G_1 und G_2 als GMO in "geeigneten" Modellen auffassen.

- ii) $u = G_3 y$ ist Projektion von y auf $QL^\perp$
 entlang $L \oplus (L^\perp \cap K(Q))$.

$$Gy = l = P_L (I - G_3) y. \quad (\text{Theorem 1(4)})$$

Es lassen sich G_3 und P_L als GMO in "geeigneten" Modellen auffassen.

Dieses Verfahren bietet sich an, falls P_L bekannt (siehe z.B. Missing-Data-Problem, Theorem 1(7)) oder leicht zu berechnen ist.

Darstellung von L als direkte Summe von Teilräumen

$$L = L_1 \oplus L_2$$

Setze $L_2 := L \cap QV$, $L_1 := L \cap L_2^\perp$ orthogonales Komplement von L_2 bezüglich eines "geeigneten" Skalarproduktes in L . Die Wahl des Skalarproduktes und damit die Wahl von L_1 hängt von der Darstellung ab, mit der L angegeben ist. Eine sehr allgemeine Darstellung ist etwa $L = AU$, wobei U ein reeller linearer Raum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, $A: U \rightarrow V$ eine lineare Abbildung, A' die Adjungierte von A bezügl. $\langle \cdot, \cdot \rangle$ ist. (siehe auch Corollar zu Theorem 1(2))

Lemma 1.1(1) Sei $(AA')^-$ eine beliebige g -Inverse zu AA' . Durch $(\cdot, (AA')^- \cdot)$ wird auf L ein Skalarprodukt definiert.

Beweis: folgt aus der Tatsache, daß $A'(AA')^-A$ orthogonale Projektion auf $A'V \subseteq U$ bezüglich $\langle \cdot, \cdot \rangle$ ist.

Lemma 1.1(2) Setze $W := (I - QQ^+)A$,
 $L_1 := AW^+WU$ und $L_2 := A(I - W^+W)U$,
 so gilt:

- i) $L_1 \subseteq L$ und $L_1 + L_2 = L$
- ii) $L_1 \perp_0 L_2$ bezüglich $(\cdot, (AA')^- \cdot)$
- iii) $L = L_1 \oplus L_2$
- iv) $L_2 = L \cap QV$ und $L_1 = AA'K(Q)$

Beweis:

- i) $L_1 \subseteq AU = L$ und $L_1 + L_2 \subseteq L$.

Sei $x = Au \in L$, so ist wegen

$$u = W^+Wu + (I - W^+W)u \quad Au \in L_1 + L_2.$$

- ii) Sei $x = AW^+Wu \in L_1$ und $z = A(I - W^+W)v \in L_2$,

so gilt:

$$\begin{aligned}(x, (AA')^{-1}z) &= \langle W^+Wu, A'(AA')^{-1}A(I-W^+W)v \rangle \\ &= \langle W^+Wu, (I-W^+W)v \rangle = 0,\end{aligned}$$

da W^+W orthogonale Projektion auf $W'U \subseteq A'U$.

iii) folgt aus i) und ii).

iv) Sei $x = A(I-W^+W)u \in L_2$, so ist
 $(I-QQ^+)x = W(I-W^+W)u = 0$, also $x \in L \cap QV$.

Sei $x = Au \in L \cap QV$, so ist
 $0 = (I-QQ^+)x = Wu$, $W^+Wu = 0$ und
 $x = A(I-W^+W)u \in L_2$.

$$L_1 = AW^+WU = AW'U = AA'K(Q).$$

Corollar: Speziell für $U=V$ und $A=A'=P_L$ ist

$$i) \quad L_1 = P_L(I-QQ^+)V = P_L(I-QQ^+)L$$

$$ii) \quad L = L_1 \oplus L_2$$

$$\begin{aligned}\text{Beweis: } i) \quad L_1 &= P_L(I-QQ^+)V \\ &= P_L[(I-QQ^+)L \oplus (L^\perp \cap K(Q))] \\ &= P_L(I-QQ^+)L\end{aligned}$$

$$ii) \quad AA' = P_L = P'_L \text{ und}$$

$$\forall x, z \in L \text{ ist } (x, P_L z) = (x, z).$$

Bemerkung: QQ^+ ist durch jede andere Darstellung des Projektors P_{QV} ersetzbar.

Schätzung des "singulären" Anteils $l_1 = G_1 y \in L_1$

Es sei $G_1: V \rightarrow L_1$ der Projektor auf L_1 entlang $QV \oplus (L^\perp \cap K(Q))$.
 G_1 hängt von dem speziell gewählten L_1 ab. In jedem Fall
jedoch kann man G_1 auffassen als GMO zum Modell $M(L_1, Q)$.

Lemma 1.1(3) Sei $L = L_1 \oplus L_2$, $G_1: V \rightarrow L_1$ der obige Projektor,
so gilt:

G_1 ist GMO zum Modell $M(L_1, Q)$.

Beweis: $G_1: V \rightarrow L_1$ idempotent und
 $G_1 Q = 0$ symmetrisch.

Mit Lemma 0(4) (Symmetriekriterium)
folgt die Behauptung.

Bemerkung: Wegen $\text{Cov}_P y = G_1 Q G_1' = 0$
ist $G_1 y = G_1 E_P y$ P-f.s. $\forall P \in \mathcal{P}$.

Lemma 1.1(4) Sei $L = AU$ wie in Lemma 1.1(1),
 $L_1 = A W^+ W U$ mit $W := (I - Q Q^+) A$,
so hat der Projektor G_1 die Darstellungen:

- i) $G_1 = A W^+ = A W^+ (I - Q Q^+)$
- ii) $G_1 = A A' [(I - Q Q^+) A A' (I - Q Q^+)]^+$

Beweis:

- i) Wegen $QV \subseteq K(W') = K(W^+)$ ist
 $W^+ = W^+ (I - Q Q^+)$.

Für $y \in QV$ ist $A W^+ (I - Q Q^+) = 0$.

Für $y \in L^\perp \cap K(Q)$ ist

$W' y = A' (I - Q Q^+) y = A' y = 0$, also $W^+ y = 0$.

Für $y = A W^+ W u \in L_1$ ist

$A W^+ y = A W^+ (I - Q Q^+) A W^+ W u = A W^+ W u = y$.

ii) Sei $B := (I - QQ^+)AA'(I - QQ^+) = B'$

Für $y \in QV$ ist $B^+y = By = 0$.

Für $y \in L^\perp \cap K(Q)$ ist

$$B^+y = By = (I - QQ^+)AA'y = 0.$$

Wegen $K(Q) = (I - QQ^+)L \oplus (L^\perp \cap K(Q))$ ist

$$B^+V = BV = (I - QQ^+)L.$$

B^+B ist orthogonale Projektion auf BV .

Für $y = AA'(I - QQ^+)l$ mit $l \in L$ (Lemma 1.1(2)) ist

$$\begin{aligned} AA'B^+y &= AA'B^+(I - QQ^+)y = AA'B^+Bl = \\ &= AA'B^+B(I - QQ^+)l = AA'(I - QQ^+)l = y. \end{aligned}$$

Corollar: Für $U=V$ und $A=A'=P_L$ gilt:

$$i) \quad G_1 = P_L W^+ = W^+(I - QQ^+) = W^+$$

$$ii) \quad G_1 = P_L [(I - QQ^+)P_L(I - QQ^+)]^+$$

$$iii) \quad G_1 = [P_L(I - QQ^+)P_L]^+(I - QQ^+)$$

Beweis:

i) und ii) folgen unmittelbar aus Lemma 1.1(4).

$$iii) \quad \text{Sei } C := P_L(I - QQ^+)P_L = C'.$$

Für $y \in QV$ ist $C^+(I - QQ^+)y = 0$.

Für $y \in L^\perp \cap K(Q)$ ist

$$C^+(I - QQ^+)y = C^+y = Cy = 0.$$

Für $y = P_L(I - QQ^+)l \in L_1$ (siehe Corollar zu Lemma 1.1(2)) ist

$$\begin{aligned} C^+(I - QQ^+)y &= C^+P_L(I - QQ^+)y = C^+C(I - QQ^+)l = \\ &= P_L C^+C(I - QQ^+)l = P_L(I - QQ^+)l = y. \end{aligned}$$

denn C^+C orthogonale Projektion auf

$$C'V = CV = P_L(I - QQ^+)L = L_1.$$

Schätzung des "regulären" Anteils $l_2 = G_2(y - l_1)$

Lemma 1.1(5) Sei $\tilde{V} := QV \oplus (L^\perp \cap K(Q))$, $G_2: \tilde{V} \rightarrow L_2$ Projektion auf $L_2 = L \cap QV$ entlang $QL^\perp \oplus (L^\perp \cap K(Q))$, so gilt:

G_2 ist GMO zum Modell $M(L_2, Q)$ in $\tilde{V}$.

Beweis: $\tilde{V} \cap L_2^\perp \subseteq L_2^\perp = (L \cap QV)^\perp = L^\perp + K(Q)$,
 $Q(\tilde{V} \cap L_2^\perp) \subseteq QL^\perp$.

Mit Lemma 0(3) (Projektionskriterium) folgt die Behauptung.

Lemma 1.1(6) Sei $L=AU$ wie in Lemma 1.1(1), $L_2 = L \cap QV$, G_2 der Projektor aus Lemma 1.1(5), so hat G_2 die Darstellung:

$$G_2 = \bar{W} [\bar{W}' Q^+ \bar{W}]^- \bar{W}' Q^+ \\ \text{mit } \bar{W} := A(I - W^+ W), \quad W := (I - QQ^+)A$$

Beweis: Wegen Lemma 1.1(2) ist
 $L_2 = A(I - W^+ W)U = \bar{W}U \subseteq QV = QV$.

Mit Corollar zu Theorem 1(2) (Aitken-Formel) folgt die Behauptung.

Corollar: Speziell für $U=V$, $A=P_L$ ergibt sich:

$$G_2 = \bar{W} [\bar{W}' Q^+ \bar{W}]^- \bar{W}' Q^+ = [\bar{W} Q^+ \bar{W}]^+ Q^+ \\ \text{mit } \bar{W} := P_L(I - W^+ W) = I - W^+ W = \bar{W}' \text{ und} \\ W := (I - QQ^+)P_L.$$

Theorem 1(3) Sei U ein reeller Vektorraum mit Skalarprodukt $\langle ., . \rangle$,

$A: U \rightarrow V$ eine lineare Abbildung, A' die Adjungierte zu A bezügl. $\langle ., . \rangle$.

Sei $L = AU$, $Q \in S_+(V)$ beliebig,

$G: V \rightarrow L$ der Projektor auf L entlang $QL^\perp \oplus (L^\perp \cap K(Q))$.

G hat dann die Darstellung:

$$\begin{aligned} G &= G_1 + \bar{W} [\bar{W}' Q^+ \bar{W}]^- \bar{W}' Q^+ (I - G_1) \\ &= G_1 + A [\bar{W}' Q^+ \bar{W}]^+ A' Q^+ (I - G_1) \\ &\quad \text{mit } \bar{W} := A(I - W^+ W), \quad W := (I - Q Q^+) A \\ &\quad G_1 \text{ gemäß Lemma 1.1(4)} \end{aligned}$$

Corollar 1 Für $L = \mathcal{U}(X) \subseteq \mathbb{R}^n$ gilt insbesondere:

$$\begin{aligned} G &= XW^+ + X [\bar{W}' Q^+ \bar{W}]^+ X' Q^+ (I - XW^+) \\ &\quad \text{mit } \bar{W} := X(I - W^+ W), \quad W := (I - Q Q^+) X \\ &\quad (\text{siehe DRYGAS 1970 (3.188),} \\ &\quad \text{"Schönfelds Formel"}) \end{aligned}$$

Corollar 2 Für einen beliebigen Teilraum $L \subseteq V$ gilt:

$$\begin{aligned} G &= G_1 + \bar{W} [\bar{W} Q^+ \bar{W}]^- \bar{W} Q^+ (I - G_1) \\ &= G_1 + [\bar{W} Q^+ \bar{W}]^+ Q^+ (I - G_1) \\ &\quad \text{mit } \bar{W} := I - W^+ W, \quad W := (I - Q Q^+) P_L \\ &\quad G_1 \text{ gemäß Corollar zu Lemma 1.1(4)} \end{aligned}$$

Schätzung von $u = G_3 y \in QL^\perp$

Lemma 1.1(7) Sei $G_3: V \rightarrow QL^\perp$ die Projektion auf $QL^\perp$ entlang $L \oplus (L^\perp \cap K(Q))$, so gilt:

G_3 ist GMO zum Modell $M(QL^\perp, P_L)$.

Beweis: $G_3: V \rightarrow QL^\perp$ ist idempotent und $G_3 P_L = 0$ symmetrisch.

Mit Lemma 0(4) (Symmetriekriterium) folgt die Behauptung.

(vergleiche auch Lemma 1.1(3))

Lemma 1.1(8) Sei $Q \in S_+(V)$ beliebig,

$\Rightarrow \exists Q^{\frac{1}{2}} \in S_+(V): Q^{\frac{1}{2}} Q^{\frac{1}{2}} = Q$

(symmetrische Quadratwurzel zu Q)

Beweis: siehe HALMOS 1958, §72

Corollar: $Q^{\frac{1}{2}}$ symmetrische Quadratwurzel zu Q ,
so gilt: $QV = Q^{\frac{1}{2}}V$.

Beweis: ich zeige: $K(Q^{\frac{1}{2}}) = K(Q)$:

$$Q^{\frac{1}{2}}x = 0 \Rightarrow Qx = Q^{\frac{1}{2}}Q^{\frac{1}{2}}x = 0.$$

$$0 = Qx = Q^{\frac{1}{2}}Q^{\frac{1}{2}}x \Rightarrow Q^{\frac{1}{2}}x \in Q^{\frac{1}{2}}V \cap K(Q^{\frac{1}{2}}) = 0.$$

Lemma 1.1(9) $QL^\perp = Q^{\frac{1}{2}}Z^+ZV$ und $(L \cap QV) = Q^{\frac{1}{2}}(I - Z^+Z)V$
mit $Z := (I - P_L)Q^{\frac{1}{2}}$.

Beweis: Wegen Lemma 1.0(5) ist

$$Q^{\frac{1}{2}}V = QV = (L \cap QV) \oplus QL^\perp$$

$$= Q^{\frac{1}{2}}(Q^{\frac{1}{2}})^+ K(P_L) \oplus (Q^{\frac{1}{2}}V \cap P_L V)$$

$(L \cap QV) \perp_0 QL^\perp$ bez. des durch $Q^- = (Q^{\frac{1}{2}}Q^{\frac{1}{2}})^-$ auf QV definierten Skalarproduktes.

Mit Lemma 1.1(2) ($U=V$, $A \rightarrow Q$, $Q \rightarrow P_L$) folgt die Behauptung.

Die Zerlegung von L in $AA'K(Q) \oplus (L \cap QV)$ und die Zerlegung von QV in $(L \cap QV) \oplus QL^\perp$ sind symmetrische (rein geometrische) Probleme. Die wahrscheinlichkeitstheoretische Bedeutung von L und Q spielen hierfür keine Rolle.

Lemma 1.1(10) Sei $G_3: V \rightarrow QL^\perp$ der Projektor aus Lemma 1.1(7), G_3 hat die Darstellungen:

$$i) \quad G_3 = Q^{\frac{1}{2}}Z^+ = Q^{\frac{1}{2}}Z^+(I-P_L)$$

$$ii) \quad G_3 = Q[(I-P_L)Q(I-P_L)]^+$$

Beweis: folgt unmittelbar aus Lemma 1.1(9) und Lemma 1.1(4).

Corollar: Falls $Q \in S_+(V)$ idempotent, so hat G_3 die weitere Darstellung:

$$G_3 = [Q(I-P_L)Q]^+(I-P_L)$$

Beweis: folgt aus Corollar zu Lemma 1.1(4).

Theorem 1(4) Sei $Q \in S_+(V)$ beliebig, sei G der Projektor aus Theorem 1(3). G hat die Darstellungen:

$$i) \quad G = P_L(I - Q^{\frac{1}{2}}Z^+) \text{ mit } Z := (I-P_L)Q^{\frac{1}{2}}$$

$$ii) \quad G = P_L(I - Q[(I-P_L)Q(I-P_L)]^+)$$

Falls $Q \in S_+(V)$ idempotent, so gilt ferner:

$$iii) \quad G = P_L(I - [Q(I-P_L)Q]^+(I-P_L)).$$

Theorem 1(3) und 1(4) liefern eine Reihe von Darstellungen desselben Projektors G , d.h. im Anwendungsfall erhält man dieselben Matrizen G .

Welche der angegebenen Darstellungen geeigneter ist, hängt vom konkreten Modell ab, im Wesentlichen von der Frage, welche der auftretenden Inversen am einfachsten zu berechnen ist.

Für andere Darstellungen des linearen Teilraumes L lassen sich Theorem 1(3) und 1(4) in naheliegender Weise verallgemeinern.

1.1.2. Majorante mit quasiregulärem und regulärem Q

Wegen Lemma 0(7) (Majorantenkriterium) ist es hinreichend, einen GMO für ein "größeres" Modell zu berechnen.

Falls es ein bezüglich L quasireguläres $T \in S_+(V)$ gibt so daß $M(L, Q) \leq M(L, T)$, so läßt sich mit Hilfe der Aitken-Formel (Corollar zu Theorem 1(2)) ein GMO $G: V \rightarrow L$ zum Modell $M(L, T)$ berechnen. G ist dann auch GMO zum Modell $M(L, Q)$.

Lemma 1.1(11) Sei $L = AU$ wie in Theorem 1(3) und 1(4),

$T := AA' + Q$, so gilt:

- i) $T \in S_+(V)$ quasiregulär bezügl. L
- ii) $M(L, Q) \leq M(L, T)$

Beweis: $TV \subseteq AA'V + QV$ und

$TV \supseteq AA'V + QV$, denn sei $x \in K(T) = (TV)^\perp$,
so ist $0 = (Tx, x) = (A'x, A'x) + (Qx, x)$
und $A'x = Qx = 0$, also $x \in L^\perp \cap K(Q) = (L + QV)^\perp$

- i) $QV \subseteq L + QV = TV$ und $QL^\perp = TL^\perp$
- ii) $L \subseteq L + QV = TV$.

Theorem 1(5) Sei $Q \in S_+(V)$ beliebig, $L = AU$ wie in Theorem 1(3) und 1(4), $T := AA' + Q$ und $G: V \rightarrow L$ der Projektor auf L entlang $QL^\perp \oplus (L^\perp \cap K(Q))$.

G hat dann die Darstellung:

$$G = A(A'T^+A)^-A'T^+$$

Beweis: Wegen Corollar zu Theorem 1(2) ist $\bar{G} := A(A'T^+A)^-A'T^+$ GMO zum Modell $M(L, T)$ und $\bar{G} = 0$ auf $K(T)$.

Wegen Lemma 0(7) ist $\bar{G}$ auch GMO
zum Modell $M(L, Q)$.

Da $K(T) = (TV)^\perp = (L + QV)^\perp = L^\perp \cap K(Q)$,
ist $\bar{G} \equiv G$.

Corollar 1 Für $L = \mathcal{M}(X) \subseteq \mathbb{R}^n$ gilt insbesondere:

$$G = X(X'T^+X)^-X'T^+ \quad \text{mit } T := XX' + Q$$

(vergleiche RAO 1973, (3.24))

Corollar 2 Für einen beliebigen Teilraum $L \subseteq V$ gilt:

$$G = P_L(P_L T^+ P_L)^- P_L T^+ = [P_L T^+ P_L]^+ T^+$$

$$\text{mit } T := P_L + Q$$

(vergleiche DRYGAS 1977, Corollar 2.2.b)

Bemerkung: Es reicht, $T \in S_+(V)$ so zu wählen, daß
 $TV = L + QV$ ist.

(vergleiche RAO 1971, Theorem 4.3., RAO 1972
(1.9))

Lemma 1.1(12) Sei $T \in S_+(V)$ so daß $TV = L + QV$,

$$Q := T + (I - TT^+),$$

so gilt:

i) $Q \in S_+(V)$ regulär

ii) $M(L, T) \subseteq M(L, Q)$.

Beweis: $T \in S_+(V)$, $(I - TT^+) \in S_+(V) \Rightarrow Q \in S_+(V)$.

$$QV = TV \oplus (I - TT^+)V = TV \oplus K(T) = V.$$

$TV \subseteq QV$ und $TL^\perp \subseteq QL^\perp$, also

$$M(L, T) \subseteq M(L, Q).$$

Bemerkung: $Q = I + T^{\frac{1}{2}}(I - T^-)T^{\frac{1}{2}}$,

$T^{\frac{1}{2}}$ symmetrische Quadratwurzel
zu T .

Theorem 1(6) Sei $Q \in S_+(V)$ beliebig,
 $L = AU$ wie in Theorem 1(3) und 1(4),
 $G: V \rightarrow L$ der Projektor auf L entlang
 $QL^\perp \oplus (L^\perp \cap K(Q))$,

$\Rightarrow \exists \tilde{Q} \in S_+(V)$ so daß $\tilde{Q}$ regulär und

$$G = A(A'\tilde{Q}^{-1}A)^{-1}A'\tilde{Q}^{-1}.$$

Insbesondere ist für beliebiges $y \in V$
 $l_0 = Gy$ die eindeutig bestimmte Lösung des
 Extremalproblems

$$(y - l, \tilde{Q}^{-1}(y - l)) \rightarrow \min_{l \in L}.$$

Beweis: Sei $\tilde{Q}$ wie in Lemma 1.1(12):

$$\begin{aligned}\tilde{Q}L^\perp &= TL^\perp \oplus (I - TT^+)L^\perp = QL^\perp \oplus (I - TT^+)L^\perp \\ &= QL^\perp \oplus (L^\perp \cap K(Q)),\end{aligned}$$

$$\begin{aligned}\text{da } (L^\perp \cap K(Q)) &= K(T) = (I - TT^+)V = \\ &= (I - TT^+)L^\perp \oplus (L \cap K(T)) = (I - TT^+)L^\perp.\end{aligned}$$

In Theorem 1(5) und 1(6) wurden jeweils "größere" Modelle zur Berechnung eines GMO für $M(L, Q)$ angegeben.

Es ist $M(L, Q) \leq M(L, T) \leq M(L, \tilde{Q})$. Die Modelle sind nicht äquivalent im Sinn von Lemma 0(5), dennoch sind die Projektoren aus Theorem 1(3) - 1(6) identisch (nicht nur P-fast sicher).

Anmerkungen zu 1.1.

- 1) Das Prinzip der "schrittweisen GM-Schätzung" wurde von Drygas entwickelt (DRYGAS 1970, Seite 108/109 und 1971, Seite 310/311):

D. betrachtet drei Modelle:

- (A) $M(L_1 + L_2, Q)$
(B) $M(L_1, Q)$
(C) $M((I - G_1)L_2, Q)$

G_1 und G_2 sind so zu bestimmen, daß G_1 GMO zu (B), $G_1 L_2 \subseteq L_1$ und G_2 GMO zu (C) ist.

D. zeigt, daß dann $G := G_1 + G_2(I - G_1)$ GMO zu (A) ist.

Am Beispiel des allgemeinen Regressionsmodells $L := \mathcal{U}(X)$ ("Schönfelds Formel") definiert D. L_1 und L_2 gerade so wie in Lemma 1.1(2) (siehe DRYGAS 1970), so daß meine Ausführungen als eine Verallgemeinerung von Schönfelds Formel auf beliebige Modelle $M(L, Q)$ anzusehen ist, wobei ich alternative Darstellungen von G_1 angegeben habe. (vergleiche auch SCHÖNFELD 1967, Theorem 1)

- 2) Die Darstellungen von G aus Theorem 1(4) i) und ii) wurden im Prinzip bereits von Albert angegeben (ALBERT 1973, (2), (3) bzw. ((4), (5))).

Eine 1(4)i) entsprechende Darstellung findet sich ferner bei Rao (RAO 1973, 3(23)).

Drygas gibt einen GMO für $M(L, Q)$ mit $L := A^{-1}(b)$ an, der für $b=0$ und $A = I - P_L$ auf eine ähnliche Darstellung führt (DRYGAS 1970, (3.79)).

Der Schwerpunkt meiner Darstellung liegt darin, die geometrische Konstruktion dieser Schätzer als eine Erweiterung des Prinzips der schrittweisen GM-Schätzung sichtbar zu machen.

- 3) Die Darstellung 1(4)iii) erscheint im Zusammenhang mit der Fragestellung dieses Abschnittes 1.1. wenig motiviert, da Q als Kovarianzoperator i.a. nicht idempotent ist.

Ich habe diese Formel speziell im Hinblick auf das Missing-Data-Problem angegeben (Abschnitt 1.2.). Dort betrachte ich formal ein Modell $M(L, I-P)$ mit $P = P' = P^2$ (siehe Lemma 1.2(5) und 1.2(6) mit Corollar 1 und 2).

- 4) Rao hat in zahlreichen Arbeiten, ausgehend von dem Gedanken einen einzelnen Schätzwert als Lösung eines Extremalproblems zu bestimmen (Verallgemeinerung der Methode der kleinsten Quadrate), gezeigt, daß Q im Falle der Singularität durch einen bezüglich L quasi-regulären Operator $T \in S_+(V)$ zu ersetzen ist, für den lediglich $TV \geq L + QV$ gelten muß.
(siehe z.B. RAO 1971 a), S. 382 ff, insbesondere Theorem 4.3; 1972, (1.9) und (1.10); 1973; MITRA & MOORE 1973)

Da Rao sich nur für die Schätzwerte G_Y interessiert, geht aus seinen Arbeiten nicht hervor, ob die jeweils berechneten Projektoren untereinander und mit den von anderen Autoren angegebenen Darstellungen identisch sind.

Ich habe mich auf Darstellungen des in Theorem 1(3) und 1(4) angegebenen Projektors beschränkt.

Das von mir eingeführte Konzept der Äquivalenz- und Ordnungsrelation zwischen linearen Modellen ermöglicht es, die Klasse der "zulässigen" $T \in S_+(V)$ so zu erweitern, daß man in jedem Fall ein reguläres $\tilde{Q} \in S_+(V)$ und somit eine Norm in V angeben kann, so daß G gerade die Orthogonalprojektion auf L bezüglich dieser Norm ist (siehe Theorem 1(6)).

1.2. Missing-Data-Problem

1.2.0. Formulierung des Problems - Das Modell $M(\mathbb{P}, \mathbb{P}\mathbb{Q}\mathbb{P})$

Sei V ein reeller linearer Raum mit Skalarprodukt $(\cdot, \cdot)$,
 $\dim V = n$.

Sei $\{e_1, \dots, e_n\}$ ein vollständiges Orthonormalsystem, d.h.
 $(e_i, e_j) = \delta_{ij}$ und $V = [e_1, \dots, e_n]$.

Dann ist für $x \in V$, $z \in V$: $x = \sum (e_i, x) e_i =: \sum x_i e_i$ und $(x, z) = \sum x_i z_i$.

Seien $y_1, \dots, y_n$ reellwertige Einzelmessungen, d.h.

$y_i: \Omega \rightarrow \mathbb{R}$ reelle Zufallsvariable.

Setze $y := \sum y_i e_i: \Omega \rightarrow V$, dann ist $y_i = (y, e_i)$.

Für y mögen die Voraussetzungen aus Kapitel 0. gelten.

Annahme: Einzelne Beobachtungen fehlen oder sind nicht
 verwertbar.

Statt $y = (y_1, \dots, y_n)'$ wird nur

$\mathbb{P}y := (\Delta_1 y_1, \dots, \Delta_n y_n)'$ beobachtet,

dabei sei $\Delta_i := \begin{cases} 0 & \text{falls } i\text{-te Beobachtung fehlt} \\ 1 & \text{falls } i\text{-te Beob. vorhanden} \end{cases}$

der Indikator für das Fehlen der i -ten Beobach-
 tung, $\mathbb{P} := \text{diag}(\Delta_1, \dots, \Delta_n) = \mathbb{P}' = \mathbb{P}^2$

Ich setze $\mathbb{P}$ als nicht-stochastischen Operator voraus.

(Eine hinreichende Bedingung wäre, daß $\mathbb{P}$ vom Beobachtungs-
 vektor y stochastisch unabhängig ist, d.h. daß die Tatsache
 des Fehlens eines Meßwertes keinen Rückschluß auf die zu
 erwartende Größe dieses Wertes zuläßt.- siehe auch

KRUSKAL 1961, Seite 440)

P bestimmt dann eine orthogonale Zerlegung von V in
 $V = PV \oplus (I-P)V =: V_1 \oplus V_2$.

Wegen $E_P Py = PE_P y \in PL$, $PL = [\{ E_P Py \mid P \in \mathcal{P} \}]$ und
 $\text{Cov}_P Py = PQP \in S_+(V)$ ist als neues Modell $M(PL, PQP)$ in PV
zu betrachten.

Lemma 1.2(1) i) $PQP(PV \cap (PL)^\perp) = PQP(PL)^\perp = PQ(L^\perp \cap PV)$

ii) $K(PQP) = (K(Q) \cap PV) \oplus V_2$

Beweis: i) $(PL)^\perp = P^{-1}L^\perp = (L^\perp \cap PV) \oplus V_2$,
 $(PL)^\perp \cap PV = L^\perp \cap PV$,

damit ist $PQP(PV \cap (PL)^\perp) = PQ(L^\perp \cap PV)$
 $= PQP(PL)^\perp$.

ii) $K(Q) \cap PV \subseteq K(PQP)$, $V_2 \subseteq K(PQP)$.

Sei $PQP x = 0$, so ist

$0 = (PQP x, x) = (QP x, Px)$, $QP x = 0$,

$Px \in K(Q) \cap PV$ und

$x = Px + (I-P)x \in (K(Q) \cap PV) \oplus V_2$.

Corollar 1 $PV = PL \oplus W_2$ mit $W_2 := PQ(L^\perp \cap PV) \oplus (L^\perp \cap PV \cap K(Q))$

Beweis: $PQP(PV \cap (PL)^\perp) = PQ(L^\perp \cap PV)$ und

$(PL)^\perp \cap K(PQP) \cap PV = L^\perp \cap PV \cap K(Q)$.

Mit Corollar 2 zu Lemma 1.0(6) folgt
die Behauptung.

Corollar 2 $V = PL \oplus W_2 \oplus V_2$.

Corollar 3 $G: PV \rightarrow PL$ Projektor auf PL entlang W_2 ,

$\rightarrow G$ ist GMO zum Modell $M(PL, PQP)$ in PV .

Beweis: folgt aus Lemma 0(3) (Projektions-
kriterium),

siehe auch Corollar 2 zu Lemma 1.0(6).

Lemma 1.2(2) Sei $\tilde{G}:V \rightarrow PL$ der Projektor auf PL entlang $W_2 \oplus V_2$, so gilt:

- i) $\tilde{G}$ ist GMO zum Modell $M(PL, PQP)$ in V .
- ii) $\forall y \in V, P \in \mathcal{P}$ ist $\tilde{G}Py$ GMS für $E_P Py \in PL$.

Beweis: i) folgt aus Lemma 0(3) mit

$$K(PQP) \cap (PL)^\perp = (L^\perp \cap PV \cap K(Q)) \oplus V_2$$

ii) siehe Corollar 3 zu Lemma 1.2(1).

Im Folgenden suche ich $\tilde{G}:V \rightarrow PL$ als Projektor auf PL entlang $W_2 \oplus V_2$. (siehe auch Bemerkung im Anschluß an Lemma 1.0(6))

$PQP \in S_+(V)$ ist auf V singulär. Mit den Formeln aus Kapitel 1.1. läßt sich $\tilde{G}$ jedoch prinzipiell immer berechnen.

1.2.1. Schätzung von $E_P Y$ aufgrund der Kenntnis von P_Y

Falls nur "wenige" Beobachtungen fehlen, d.h. falls $P = P' = P^2$ so daß $\dim PL = \dim L$ (also $PL \cong L$), so kann man wegen $E_P \tilde{G} P_Y = P E_P Y =: P l = E_P Y \in L$ aufgrund der Kenntnis von P_Y erwartungstreu schätzen:

Lemma 1.2(3) Sei $Q \in S_+(V)$ beliebig,

$W_2 = W_2(Q)$ wie in Corollar 1 zu Lemma 1.2(1),
 $P = P' = P^2$ so daß $\dim PL = \dim L$, so gilt:

- i) $L \cap (W_2 \oplus V_2) = 0$
- ii) $L \cap W_2 = L \cap V_2 = 0$

Beweis: i) Wegen $L \cap (W_2 \oplus V_2) \subseteq L$, P regulär auf L ist

$$L \cap (W_2 \oplus V_2) = P^{-1} P [L \cap (W_2 \oplus V_2)] = P^{-1} (PL \cap W_2) = 0.$$

Corollar $V = L \oplus W_2 \oplus V_2$

Beweis: Wegen Corollar 1 zu Lemma 1.2(1) ist $\dim V = \dim PL + \dim (W_2 \oplus V_2)$.

Wegen $\dim PL = \dim L$ folgt mit Lemma 1.2(3) die Behauptung.

Bemerkung: Sei für $x \in V$ $x = Pl + u + v$ die eindeutig bestimmte Zerlegung mit $l \in L$, $u \in W_2$ und $v \in V_2$, dann ist $x = l + u + (I-P)(v-l)$ die eindeutig bestimmte Zerlegung gemäß $V = L \oplus W_2 \oplus V_2$.

Lemma 1.2(4) Sei $Q \in S_+(V)$ beliebig,

$P = P' = P^2$ so daß $\dim PL = \dim L$,

$\tilde{G}: V \rightarrow PL$ der Projektor auf PL entlang $W_2 \oplus V_2$,

$G_0: V \rightarrow L$ der Projektor auf L entlang $W_2 \oplus V_2$,

so gilt:

$$i) \quad PG_0 = \tilde{G}$$

$$ii) \quad E_P G_0 P Y = E_P G_0 Y = E_P Y \quad \forall Y \in V, \forall P \in \mathcal{P}.$$

Beweis: i) Wegen $G_0(I-P) = 0$ ist $G_0 = G_0 P$
und $PG_0: V \rightarrow PL$ idempotent.

$$K(PG_0) = K(\tilde{G}) = W_2 \oplus V_2, \text{ also } PG_0 = \tilde{G}.$$

$$ii) \quad E_P G_0 Y = G_0 E_P Y = E_P Y \in L, \\ \text{da } G_0 \text{ idempotent.}$$

Bemerkung: 1) $\forall Y \in V, \forall P \in \mathcal{P}$ ist $G_0 P Y$ erwartungstreuer
Schätzer für $E_P Y \in L$.

2) G_0 ist in dem Sinn optimal, daß PG_0 GMO
zum Modell $M(PL, PQP)$ in V ,

$$\text{bzw. } \text{Var}_P(x, G_0 Y) \quad \text{minimal} \\ \forall P \in \mathcal{P} \text{ und } \forall x \in PL.$$

Lemma 1.2(5) Sei $\dim PL = \dim L$,

$G_0: V \rightarrow L$ wie oben, so gilt:

G_0 ist GMO zum Modell $M(L, (I-P))$.

Beweis: $G_0: V \rightarrow L$ idempotent,

$$G_0(I-P) = 0 \text{ symmetrisch.}$$

Mit Lemma 0(4) (Symmetriekriterium)
folgt die Behauptung.

Bemerkung: G_0 ist für beliebiges $Q \in S_+(V)$ i.a. nicht der
GMO aus Kapitel 1.1. (siehe unten)

1.2.2. "Ersetzen fehlender Daten" - Der Fall $Q = I$

Lemma 1.2(6) $V_2 = (I-P)L^\perp$

$$\begin{aligned} \text{Beweis: } V_2 &= (I-P)V = (I-P)L^\perp \oplus (V_2 \cap L) \\ &= (I-P)L^\perp \text{ (siehe Lemma 1.2(3))} \end{aligned}$$

Corollar 1 Für $Q = I$, $G_0: V \rightarrow L$ Projektor auf L entlang $W_2 \oplus V_2$ gilt:

G_0 ist der GMO aus Kapitel 1.1. zum Modell $M(L, (I-P))$.

$$\begin{aligned} \text{Beweis: } W_2(I) &= L^\perp \cap PV, \\ V &= L \oplus (L^\perp \cap PV) \oplus (I-P)L^\perp \\ &= [L \oplus (I-P)L^\perp] \oplus (L^\perp \cap PV) \end{aligned}$$

Mit Corollar 2 zu Lemma 1.0(6)
folgt die Behauptung.

Corollar 2

- i) $G_0 = (PP_L)^+ = P_L(PP_L P)^+ = (P_L PP_L)^+ P$
- ii) $G_0 = P_L(I - [(I-P)(I-P_L)(I-P)]^+(I-P_L))$
- iii) $G_0 = P_L(P_L T^+ P_L)^- P_L T^+$ mit $T := P_L + (I-P)$

Beweis:

i) folgt aus Theorem 1(3), bzw. Lemma 1.1(4)
mit $A := P_L$, $W := PP_L$

ii) folgt aus Theorem 1(4) iii)

$G_{03} := [(I-P)(I-P_L)(I-P)]^+(I-P_L)$ ist
Projektion auf $(I-P)L^\perp = V_2$ entlang
 $L \oplus (L^\perp \cap PV)$.

iii) folgt aus Theorem 1(5)

Damit sind verschiedene Möglichkeiten gegeben, G_0 bzw. $\tilde{G}$ mit Hilfe des für das vollständige Modell $M(L, I)$ bekannten Projektors P_L zu berechnen.

Die Darstellung ii) entspricht dem üblichen Verfahren, fehlende Daten durch Schätzwerte zu ersetzen:

Lemma 1.2(7) Sei $\dim PL = \dim L$,

$x_0 := APy := -[(I-P)(I-P_L)(I-P)]^+(I-P_L)Py$,
so gilt:

i) $\forall y \in V$ ist x_0 die eindeutig bestimmte Lösung des linearen Gleichungssystems

$$(*) \quad x = (I-P)P_L(Py + x)$$

ii) $\forall y \in V$ ist x_0 die eindeutig bestimmte Lösung des Extremalproblems

$$\|Py + x - P_L(Py + x)\| \rightarrow \min_{x \in V_2}.$$

Beweis:

i) Sei $B := (I-P)(I-P_L)(I-P) = B'$,
 $BV = (I-P)L^\perp = V_2$, denn $L^\perp = (I-P_L)V_2 \oplus (L^\perp \cap PV)$.
Also ist $BB^+ = (I-P)$ orthogonale Projektion auf $BV = V_2$.

Das Gleichungssystem $(*)$ ist äquivalent zu: " $Bx = (I-P)P_LPy$ und $x \in V_2$ ".

$$\begin{aligned} \text{Es ist } Bx_0 &= -BB^+(I-P_L)Py = -(I-P)(I-P_L)Py \\ &= (I-P)P_LPy \quad \text{und} \quad x_0 \in V_2. \end{aligned}$$

Die Lösung ist eindeutig, denn sei $x \in V_2$ und $Bx = 0$, so folgt:

$$x \in V_2 \cap K(B) = V_2 \cap PV = 0.$$

$$\begin{aligned} \text{ii) } \|Py + x - P_L(Py + x)\| &= \\ \|Py - PP_L(Py + x)\| + \|x - (I-P)P_L(Py + x)\| \\ \forall x \in V_2. \end{aligned}$$

Aus i) folgt unmittelbar, daß $x_0 = APy$
Lösung des Extremalproblems ist, da nach
Lemma 1.2(4)

$$\|Py - PP_L(Py + x_0)\| = \|Py - GPy\| = \min_{l \in L} \|Py - Pl\|.$$

Corollar 1 $\forall y \in V, \forall p \in P$ ist $x_0 := APy$ erwartungstreuer
Schätzer für $E_P(I-P)y$.

$$\text{Beweis: } E_P APy = (I-P)E_P G_0 Py = (I-P)E_P y = E_P(I-P)y.$$

Corollar 2 $\forall y \in V$ ist $x_0 := APy$ das eindeutig bestimmte $x \in V_2$
so daß

$$\text{i) } \tilde{G}Py = P_{PL}Py = PP_L(Py + x) \text{ und}$$

$$\text{ii) } G_0 Py = \tilde{G}Py + x.$$

$$\text{Beweis: Wegen } (PL)^\perp = (L^\perp \cap PV) \oplus V_2 \text{ ist } \tilde{G} = P_{PL}.$$

$$GP_y = PG_0 Py = PP_L(I+A)Py,$$

$$\begin{aligned} \tilde{G}Py + x_0 &= PP_L(Py + x_0) + x_0 = \\ &= P_L(Py + x_0) = G_0 Py. \end{aligned}$$

Sei $x \in V_2$ so daß i) und ii) erfüllt
sind.

$$x = (G_0 - \tilde{G})Py = (I-P)G_0 Py, \text{ d.h.}$$

x ist Lösung von (*), also eindeutig
bestimmt.

(vergleiche hierzu WILKINSON 1958 und
KRUSKAL 1961, Seite 440 ff)

Theorem 1(7) Sei $P = P' = P^2$ so daß $\dim PL = \dim L$,
 $G_o: V \rightarrow V_2$ der Projektor auf V_2 entlang
 $L \oplus (L^\perp \cap PV)$,

so gilt:

- i) $\forall y \in V, P \in \mathcal{P}$ ist $x_o := APy := -G_o Py$
 erwartungstreuer Schätzer für $E_P(I-P)y$;

$$G_o = [(I-P)(I-P_L)(I-P)]^+ (I-P_L).$$

$\forall y \in V$ ist x_o die eindeutig bestimmte
 Lösung des linearen Gleichungssystems

$$(*) \quad x = (I-P)P_L(Py + x).$$

- ii) $\forall y \in V, P \in \mathcal{P}$ ist $G_o Py := P_{PL}Py + x_o$
 erwartungstreuer Schätzer für $E_P y \in L$;

$$\forall y \in V \text{ ist } G_o Py = P_L(Py + x_o),$$

$$\text{d.h. } G_o = P_L(I + A)P.$$

- iii) $\forall y \in V, P \in \mathcal{P}$ ist $P_{PL}Py = PP_L(Py + x_o)$,

$$\text{d.h. } P_{PL} = PP_L(I + A)P.$$

1.2.3. "Ersetzen fehlender Daten" - $Q \in S_+(V)$ beliebig

Für beliebiges $Q \in S_+(V)$ sei

$W_2 = W_2(Q)$ wie in Corollar 1 zu Lemma 1.2(1),

$G_o: V \rightarrow L$ der Projektor auf L entlang $W_2 \oplus V_2$ (Lemma 1.2(4)).

Ferner sei

$\tilde{G}: V \rightarrow PL$ der Projektor auf PL entlang $W_2 \oplus V_2$,

d.h. $\tilde{G}$ GMO zum Modell $M(PL, PQP)$ in V ;

$G: V \rightarrow L$ der Projektor auf L entlang $W_1 := QL^\perp \oplus (L^\perp \cap K(Q))$;

d.h. G GMO zum Modell $M(L, Q)$.

Falls $PL \cong L$, gibt es genau ein $x \in V_2$ so daß $\tilde{G}Py + x = G_oPy$ und zwar $x_o := (G_o - \tilde{G})Py = (I - P)G_oPy$. Wegen Lemma 1.2(4) ist x_o erwartungstreue Schätzung für $E_P(I - P)y$. (vergleiche Corollar 2 zu Lemma 1.2(7))

Lemma 1.2(8) Sei $P = P' = P^2$ so daß $PL \cong L$, $Q \in S_+(V)$ beliebig,
 $x_o := (I - P)G_oPy$, so gilt:

$$x_o = -G_{o3}Py, \quad G_{o3}: V \rightarrow V_2 \text{ Projektor auf } V_2 \text{ entlang } L \oplus W_2$$

(vergleiche Corollar 2.ii) zu Lemma 1.2(6))

Beweis: Sei $y \in V$ beliebig,

$$Py = Pl + w = l + w - (I - P)l \text{ mit } l \in L, w \in W_2.$$

$$G_{o3}Py = -(I - P)l = -x_o.$$

Lemma 1.2(9) Sei $P = P' = P^2$ so daß $PL \cong L$,
 $\forall y \in V$ sei $x_o := APy := -G_{o3}Py$, so gilt:

$$G_oPy = G(Py + x_o) \iff W_2 \subseteq W_1$$

Beweis: Sei $y \in V$ beliebig,

$$Py = Pl + w = l + w - (I-P)l \quad \text{mit } l \in L, w \in W_2.$$

$$G_0 Py = l, \quad x_0 = (I-P)l,$$

$$G(Py + x_0) = G(l + w) = l + Gw.$$

$$G_0 = G(I+A)P \quad \langle \Rightarrow \rangle \quad W_2 \subseteq K(G) = W_1.$$

Corollar: $\forall y \in V$ sei $x_0 := APy := -G_0 Py$, so gilt:

$$\tilde{G}Py = PG(Py + x_0) \quad \langle \Rightarrow \rangle \quad W_2 \subseteq W_1.$$

Damit ist eine notwendige und hinreichende Bedingung dafür gegeben, daß - im Falle $\dim PL = \dim L$ - fehlende Daten "durch Schätzungen ersetzt" werden können.

Lemma 1.2(10) Sei $Pl \in L$, $W_2 \subseteq W_1$, x_0 wie oben, so gilt:

x_0 ist Lösung des linearen Gleichungssystems

$$(*) \quad x = (I-P)G(Py + x)$$

Beweis: $W_2 \subseteq W_1$, so ist $G_0 Py = G(Py + x_0)$,

$$x_0 = (I-P)G_0 Py = (I-P)G(Py + x_0).$$

Lemma 1.2(11) (*) hat höchstens eine Lösung

$$\langle \Rightarrow \rangle \quad \text{Rg}(I-P)(I-G) = \dim V_2.$$

Beweis: Sei $B := (I-P)(I-G)(I-P)$,

(*) ist äquivalent zu

$$Bx = (I-P)GPy \quad \text{und } x \in V_2.$$

(*) hat höchstens eine Lösung

$$\langle \Rightarrow \rangle \quad K(B) \cap V_2 = 0$$

$$K(B) \cap V_2 = \overline{L \oplus (W_1 \cap PV)} \cap V_2, \quad \text{denn}$$

$$\begin{aligned} x \in V_2 \cap [L \oplus (W_1 \cap PV)] &\implies Bx = (I-P)(I-G)x = 0; \\ x \in V_2 \text{ und } Bx = 0 &\implies (I-G)(I-P)x \in W_1 \cap PV, \\ x = (I-P)x \in [L \oplus (W_1 \cap PV)] \cap V_2. \end{aligned}$$

Da $V = L \oplus W_1 = L \oplus [(W_1 \cap PV) \oplus P_{W_1}V_2]$ und $\dim P_{W_1}V_2 \leq \dim V_2$, gilt:

$$\begin{aligned} K(B) \cap V_2 = 0 &\iff V = L \oplus (W_1 \cap PV) \oplus V_2, \\ K(B) \cap V_2 = 0 &\iff \dim P_{W_1}V_2 = \dim V_2. \end{aligned}$$

$\dim P_{W_1}V_2 = \operatorname{Rg} P_{W_1}(I-P) = \operatorname{Rg} (I-P)P_{W_1}$, da P_{W_1} symmetrisch, $(I-P)$ orthogonale Projektion, $\operatorname{Rg} (I-P)P_{W_1} = \dim (I-P)W_1 = \dim (I-P)(I-G)V = \operatorname{Rg} (I-P)(I-G)$.

$$K(B) \cap V_2 = 0 \iff \operatorname{Rg} (I-P)(I-G) = \dim V_2.$$

Corollar 1

- i) $\operatorname{Rg} (I-P)(I-G) = \dim V_2$
 $\iff V = L \oplus (W_1 \cap PV) \oplus V_2$
- ii) Falls $PL \cong L$, so gilt:
 $\operatorname{Rg} (I-P)(I-G) = \dim V_2 \iff PV = PL \oplus (W_1 \cap PV).$

Beweis: i) siehe Beweis zu Lemma 1,2(11)

ii) " $\Leftarrow$ " Sei $Pl = w \in PL \cap W_1$, so ist
 $l = w + (I-P)l \in L \cap [(W_1 \cap PV) \oplus V_2] = 0$,
 also $PL \cap W_1 = 0$.

" $\Rightarrow$ " Sei $l = w + v \in L$, $w \in (W_1 \cap PV)$, $v \in V_2$,
 so ist $Pl = w \in PL \cap W_1 = 0$, also $l = 0$,
 $V = L \oplus (W_1 \cap PV) \oplus V_2$.

Mit i) folgt die Behauptung.

Corollar 2 Sei $PL \equiv L$, $W_2 \subseteq W_1$, so gilt:

$\forall y \in V$ ist $x_0 := -G_0 P y$ die eindeutig bestimmte Lösung des linearen Gleichungssystems (*)

$$\langle \Rightarrow \rangle \quad W_2 = W_1 \cap PV$$

Lemma 1.2(12) Sei $PL \equiv L$, $Rg (I-P) (I-G) = \dim V_2$,

$H: V \rightarrow V_2$ der Projektor auf V_2 entlang $L \oplus (W_1 \cap PV)$, so gilt:

- i) $H = B^+ (I-G)$ mit $B := (I-P) (I-G) (I-P)$
- ii) $\forall y \in V$ ist $\bar{x} := -B^+ (I-G) P y$ die eindeutig bestimmte Lösung von (*)

Beweis:

- i) Wegen $B^+ = B^+ (I-P) = (I-P) B^+$ ist $B^+ (I-G)$ idempotent.

Für $l \in L$ ist $B^+ (I-G) l = 0$.

Für $w \in (W_1 \cap PV)$ ist $B^+ (I-P) (I-G) w = 0$.

Wegen $Rg (I-P) (I-G) = \dim V_2$ ist $K(B) \cap V_2 = 0$ (siehe Beweis zu Lemma 1.2(11)),

$$V_2 = (V_2 \cap K(B)) \oplus (I-P) B' V = (I-P) B' V = B' V,$$

$B^+ B = (I-P)$ orthogonale Projektion auf $B' V$.

Für $v \in V_2$ ist $B^+ (I-G) v = B^+ B v = v$.

- ii) $\bar{x}$ ist Lösung von (*), denn $\bar{x} \in V_2$ und $B \bar{x} = -B B^+ (I-P) (I-G) P y$.

Sei $P y = P l + w = l + w - (I-P) l$ mit $l \in L$ und $w \in (W_1 \cap PV)$, so ist $(I-P) (I-G) P y = -B l$.

$$B \bar{x} = B B^+ B l = B l = - (I-P) (I-G) P y = (I-P) G P y$$

Damit ist $\bar{x}$ Lösung von (*) (siehe Beweis zu Lemma 1.2(11)).

Die Eindeutigkeit folgt aus Lemma 1.2(11).

Lemma 1.2(13) Sei $PL \cong L$, $W_2 \subseteq W_1$, so gilt:

$$\forall_{Y \in V} \text{ ist } x_0 = B^+(I-G)PY \iff W_2 = W_1 \cap PV$$

Beweis:

" $\Leftarrow$ " Sei $W_2 = W_1 \cap PV$, so ist $\forall_{Y \in V} x_0 := -G_0 P Y$ die eindeutig bestimmte Lösung von (*), (siehe Corollar 2 zu Lemma 1.2(11)).

Wegen Corollar 1,ii) zu Lemma 1.2(11)

$$\text{ist } \text{Rg } (I-P)(I-G) = \dim V_2.$$

Wegen Lemma 1.2(12) ist dann $\bar{x}$ Lösung von (*).

$$\text{Also } x_0 = \bar{x} \quad \forall_{Y \in V}.$$

" $\Rightarrow$ " Da $W_2 \subseteq W_1$ nach Voraussetzung ist nur zu zeigen: $W_2 \cong W_1 \cap PV$,
bzw. äquivalent: $PV = PL \oplus (W_1 \cap PV)$.

Sei $Pl = w \in W_1 \cap PL$, so ist $-G_0 Pl = (I-P)l$

$$\text{und } B^+(I-G)w = 0,$$

$$\text{wegen } B^+(I-G) = G_0, \text{ also } (I-P)l = 0.$$

$$l = Pl = w \in L \cap W_1 = 0.$$

Also ist $W_1 \cap PL = 0$ und

$$PL \oplus (W_1 \cap PV) \subseteq PV = PL \oplus W_2 \subseteq PL \oplus (W_1 \cap PV).$$

Damit ist $W_2 \cong W_1 \cap PV$.

Theorem 1(8) Sei $P = P' = P^2$ so daß $\dim PL = \dim L$,
 $Q \in S_+(V)$ beliebig.

Sei ferner

$\tilde{G}: V \rightarrow PL$ der Projektor auf PL entlang $W_2 \oplus V_2$,

$G: V \rightarrow L$ der Projektor auf L entlang W_1 ,

$G_0: V \rightarrow V_2$ der Projektor auf V_2 entlang $L \oplus W_2$,

so gilt:

- i) $\forall y \in V, P \in \mathcal{P}$ ist $x_0 := APy := -G_0Py$
 erwartungstreuer Schätzer für $E_P(I-P)y$.
- ii) $\forall y \in V, P \in \mathcal{P}$ ist $G_0Py := \tilde{G}Py + x_0$
 erwartungstreuer Schätzer für $E_P y \in L$.
- iii) Folgende Aussagen sind paarweise
 äquivalent:
 - a) $W_2 \subseteq W_1$
 - b) $\forall y \in V$ ist $\tilde{G}Py = PG(Py + x_0)$, d.h.
 $\tilde{G} = PG(I + A)P$
 - c) $\forall y \in V$ ist $G_0Py = G(Py + x_0)$, d.h.
 $G_0 = G(I + A)P$
- iv) Falls $W_2 \subseteq W_1$, so sind folgende
 Aussagen paarweise äquivalent:
 - a) $W_2 = W_1 \cap PV$
 - b) $\operatorname{Rg} (I-P)(I-G) = \dim V_2$
 - c) $\forall y \in V$ ist
 $x_0 = -[(I-P)(I-G)(I-P)]^+(I-G)Py$
 - d) $\forall y \in V$ ist x_0 die eindeutig bestimmte
 Lösung des linearen Gleichungssystems
 $(*) \quad x = (I-P)G(Py + x)$

(vergleiche Theorem 1(7))

Ergänzungen zu 1.2.

- 1) In Theorem 1(6) wurde gezeigt, daß es zu beliebigem $Q \in S_+(V)$ ein reguläres $\tilde{Q} \in S_+(V)$ gibt, so daß $W_1(Q) := QL^\perp \oplus (L^\perp \cap K(Q)) = \tilde{Q}L^\perp =: W_1(\tilde{Q})$ gilt, wobei $\tilde{Q}L^\perp = L^{\perp_0}$ das orthogonale Komplement von L bezüglich des durch $(\cdot, \cdot)_0 := (\cdot, \tilde{Q}^{-1} \cdot)$ definierten Skalarproduktes ist (siehe Lemma 1.0(2)).

$P := \text{diag}(\Delta_1, \dots, \Delta_n)$ ist nach Definition symmetrisch bezüglich $(\cdot, \cdot)_0$ (siehe Abschnitt 1.2.0.).

Falls P zusätzlich symmetrisch bezüglich $(\cdot, \cdot)_0$ ist - äquivalent: $P\tilde{Q}^{-1} = \tilde{Q}^{-1}P$ oder $P\tilde{Q} = \tilde{Q}P$ -, so ist $G_0 := V \rightarrow L$ der GMO zum Modell $M(L, I-P)$ in $(V, (\cdot, \cdot)_0)$ wie in 1.1. (siehe Corollar 1 zu Lemma 1.2(6)), denn mit Lemma 1.2(6) folgt:

$$(I-P)L^{\perp_0} = \tilde{Q}(I-P)L^\perp = \tilde{Q}(I-P)V = (I-P)\tilde{Q}V = (I-P)V = V_2.$$

Mit Corollar 2 zu Lemma 1.2(6) gilt dann:

$$G_0 = P_L^\circ (I - [(I-P)(I-P_L^\circ)(I-P)]^+ (I-P_L^\circ)),$$

wobei P_L° die orthogonale Projektion auf L bezüglich $(\cdot, \cdot)_0$ sei - z.B. $P_L^\circ = P_L(P_L\tilde{Q}^{-1}P_L)^- P_L^\circ \tilde{Q}^{-1}$.

In diesem Fall - $P\tilde{Q} = \tilde{Q}P$ - ist also die Bedingung iv) aus Theorem 1(8) " $W_2 = W_1 \cap PV$ " sicher erfüllt.

(Man kann auch direkt zeigen:

$$P\tilde{Q} = \tilde{Q}P \implies W_2(Q) = W_2(\tilde{Q}) = W_1(\tilde{Q}) \cap PV = W_1(Q) \cap PV.)$$

Der Übergang zum Skalarprodukt $(\cdot, \cdot)_0$ entspricht in der üblichen Terminologie einer Transformation des Beobachtungsvektors $y \rightarrow Q^{-\frac{1}{2}}y$ bzw. einem Wechsel der Basis $\{e_1, \dots, e_n\}$ (siehe Abschnitt 1.2.0.)

Ein hinreichendes und leicht nachzuprüfendes Kriterium dafür daß fehlende Daten "durch Schätzwerte ersetzbar" sind liefert auch das folgende Lemma:

Lemma 1.2(14) $PQ = QP \implies W_2 \subseteq W_1$

$$\begin{aligned} \text{Beweis: } W_2 &= PQ(L^\perp \cap PV) \oplus (L^\perp \cap PV \cap K(Q)) \\ &= Q(L^\perp \cap PV) \oplus (L^\perp \cap PV \cap K(Q)) \\ &\subseteq QL^\perp \oplus (L^\perp \cap K(Q)) = W_1 \end{aligned}$$

Corollar: $PQ = QP$ und $Rg(I-P)(I-G) = Rg(I-P)$

$$\implies W_2 = W_1 \cap PV$$

- 2) Falls $\dim PL = \dim L$ und $Rg(I-P)(I-G) = Rg(I-P)$, so ist $V = L \oplus (W_1 \cap PV) \oplus V_2$ und $PV = PL \oplus (W_1 \cap PV)$ (siehe Corollar 1 zu Lemma 1.2(11)).

$\bar{x} := -[(I-P)(I-G)(I-P)]^+(I-G)Py$ ist dann die Projektion von Py auf V_2 entlang $L \oplus (W_1 \cap PV)$ (Lemma 1.2(12)) und $\bar{G}Py := PG(Py + \bar{x})$ ist die Projektion von Py auf PL entlang $W_1 \cap PV$.

(Denn sei $Py = Pl + w = l + w + (I-P)l$ mit $l \in L$ und $w \in W_1 \cap PV$, so ist $\bar{x} = (I-P)l$ und $\bar{G}Py = PG(Py + \bar{x}) = PG(l + w) = Pl$.)

Theorem 1(8) besagt dann gerade, daß $\bar{G} = \tilde{G} \iff W_2 \subseteq W_1$. Wegen Lemma 0(4) ist $\bar{G}$ GMO zum Modell $M(PL, PQP)$, genau dann wenn $\bar{G}PQP$ symmetrisch ist.

Wegen Lemma 0(3) ist dies äquivalent dazu daß

$PQP(PL)^\perp \subseteq W_1 \cap PV$, also $PQ(L^\perp \cap PV) \subseteq W_1 \cap PV$. Das ist aber nichts anderes als die Bedingung " $W_2 \subseteq W_1 \cap PV$ ".

Lemma 1.2(15) Sei $\dim \mathbb{P}L = \dim L$, $\text{Rg}(I-P)(I-G) = \text{Rg}(I-P)$,

$$\bar{G} := PG(I - [(I-P)(I-G)(I-P)]^+ (I-G))P,$$

so gilt:

$$\bar{G} = \bar{G} \iff \bar{G}PQP \text{ symmetrisch}$$

Zur Berechnung:

Unter den obigen Voraussetzungen ist $\bar{X}$ die eindeutig bestimmte Lösung des linearen Gleichungssystems (*) aus Theorem 1(8) (siehe Lemma 1.2(12)).

Wegen $\text{Rg}(I-P)(I-G) = \text{Rg}(I-P)$ besteht $(I-P)(I-G)(I-P) := B$ aus einer regulären Untermatrix C mit $\text{Rg}(C) = \text{Rg}(B) = \dim V_2$ und lauter Nullen (nach Definition von P), so daß nur eine Matrix vom Rang $\dim V_2$ zu invertieren ist. (siehe auch Beispiel 1.3.1.)

- 3) Ist insbesondere $QL \subseteq L$, so ist $G = P_L$ und $W_1 = L^\perp$, also $\text{Rg}(I-P)(I-P_L) = \dim (I-P)L^\perp = \dim V_2$ (Lemma 1.2(6)) und $\bar{G} = P_{\mathbb{P}L}$ (vergleiche Theorem 1(7)).

Die Frage, ob - im Falle $\dim \mathbb{P}L = \dim L$ - fehlende Daten "durch Schätzwerte ersetzt" werden können, d.h. ob $\bar{G} = P_{\mathbb{P}L}$ ist, hat folgende alternativen Formulierungen:

- i) Ist $P_{\mathbb{P}L}$ GMO zum Modell $M(\mathbb{P}L, PQP)$?
- ii) Ist $M(\mathbb{P}L, PQP) \leq M(\mathbb{P}L, P)$?
- iii) Existiert ein GMO zum Modell $M(\mathbb{P}L, P \oplus P)$, falls $\exists := \{Q, I\}$?

In diesem Fall kann man also das Missing-Data-Problem als ein spezielles Existenzproblem für gleichmäßig beste Schätzer ansehen (siehe Kapitel 2.).

Anmerkungen zu 1.2.

- 1) Die ersten Veröffentlichungen, die sich mit dem Ersetzen fehlender Daten durch Schätzwerte befassen, stammen von ALLAN & WISHART (1930) und YATES (1933).

In den darauffolgenden Jahren erschienen zahlreiche Arbeiten, die explizite Schätzformeln für Missing-Data für konkrete Versuchspläne angeben (siehe z.B. ANDERSON 1946, DE LURY 1946, THOMPSON 1956, BIGGERS 1958, etc.). Dabei werden zur Berechnung der Schätzfunktionen jeweils die Fehlerquadratsummen minimiert. Implizit wird also immer ein Modell der Form $M(L, \sigma^2 I)$ zugrunde gelegt, insbesondere wird vorausgesetzt, daß alle Einzelmessungen unkorreliert sind.

Eine allgemeine Beschreibung dieses Verfahrens findet sich bei WILKINSON (1958). Kruskal gibt eine koordinatenfreie Begründung dieser Methode an (KRUSKAL 1961) sowie eine notwendige und hinreichende Bedingung an die Zahl und Lokalisation der fehlenden Daten ($\dim PL = \dim L$).

Allen diesen Arbeiten ist gemeinsam, daß sie für die praktische Anwendung zwei nacheinander durchzuführende Operationen vorsehen:

- i) Berechnung von Schätzwerten für die fehlenden Daten, bzw. deren Erwartungswert, und
- ii) Auffüllen der Restdaten mit diesen Schätzwerten und Anwendung des bekannten Projektors.

Insbesondere werden nur einzelne Schätzwerte berechnet, nicht jedoch eine explizite Darstellung des Projektors für das reduzierte Modell angegeben.

- 2) Mir sind keine Arbeiten bekannt, die sich mit dem Problem des Ersetzens fehlender Daten unter einer allgemeineren Annahme über die Kovarianzstruktur der Beobachtungswerte befassen.

DRAPER & STONEMAN (1964) bemerken lediglich, daß bei stark korrelierten Meßwerten die üblichen Verfahren des Ersetzens fehlender Daten zu unbefriedigenden Ergebnissen führen und daher nicht verwendet werden sollten.

Offensichtlich läßt sich dieses Problem nicht in jedem Fall durch eine Transformation des Beobachtungsvektors auf die Situation " $Q=\sigma^2 I$ " zurückführen (siehe Absatz 1) der Ergänzungen zu 1.2.)

Die Bedeutung diese Problems ist vor allem im Hinblick auf die Frage der Existenz und Berechnung gleichmäßig bester Schätzer im Falle unvollständiger Daten zu sehen (siehe Abschnitt 2.2., Beispiele 1.3. und 2.3., sowie Absatz 3) der Ergänzungen zu 1.2.)

1.3. Beispiele

1.3.0. Regressionsmodell mit festen und zufälligen Effekten

Sei $V := \mathbb{R}^n$, $L := \mathcal{M}(X_1)$, wobei X_1 bekannte $n \times p$ -Matrix
 $Q := X_2 \Sigma X_2' + \sigma^2 I$ bekannt
 $(X_2, n \times q\text{-Matrix}, \Sigma = \Sigma' \geq 0 \text{ } q \times q\text{-Matrix}, \sigma^2 \geq 0)$

Äquivalente Formulierung:

$y = X_1 \beta + X_2 b + e$, β unbekannter Parameter,
 $E_P b = E_P e = 0$ und $\text{Cov}_P(b, e) = \text{diag}(\Sigma, \sigma^2 I_n)$.

Behauptung: $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$, $P = P' = P^2$ beliebig, so gilt:

- i) $P_L = X_1 (X_1' X_1)^{-1} X_1'$ ist GMO zum Modell $M(L, Q)$
- ii) $P_{PL} = P X_1 (X_1' P X_1)^{-1} X_1' P$ ist GMO zum Modell $M(PL, PQP)$.

Beweis: $X_2 \Sigma X_2' V \subseteq \mathcal{M}(X_2) \subseteq \mathcal{M}(X_1) = L$ und
 $P X_2 \Sigma X_2' P V \subseteq P \mathcal{M}(X_2) \subseteq P \mathcal{M}(X_1) = PL$, also ist
 $QL = (X_2 \Sigma X_2' + \sigma^2 I) L \subseteq L$ und $PQPL \subseteq PL$.
 Mit Corollar zu Lemma 0(4) folgt die Behauptung.

Folge: Falls $P = P' = P^2$ so daß $\text{Rg}(P X_1) = \text{Rg}(X_1)$ - d.h.
 $\dim PL = \dim L$ -, so können die fehlenden Daten
 $(I - P)y$ durch Schätzwerte $x_0 = A P y$ ersetzt werden.
 Es ist dann $P_{PL} P y = P P_L (P y + x_0)$ und
 $P_L (P y + x_0) = P_{PL} P y + x_0$ erwartungstreue Schätzung
 für $E_P y$.

1.3.1. $\mathcal{M}(X_1) = \mathcal{M}(X_2)$: Fehlende Daten sind "durch Schätzwerte ersetzbar"

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

$$\text{Sei } V := R^{nm}, \quad L := \mathcal{M}(X_1) \text{ mit } X_1 := \begin{pmatrix} 1 & A_1 & I_m \\ \vdots & \vdots & \vdots \\ 1 & A_n & I_m \end{pmatrix}, \quad A_i := \begin{pmatrix} 0 \dots 1 \dots 0 \\ \vdots & \ddots & \vdots \\ 0 \dots 1 \dots 0 \end{pmatrix}$$

$\underbrace{\hspace{10em}}_i$

A_i $n \times m$ -Matrix

$$Q := X_2 \Sigma X_2' + \sigma_e^2 I_{nm} \text{ mit } \Sigma := \text{diag}(\sigma_a^2 I_n, \sigma_b^2 I_m), \quad X_2 := \begin{pmatrix} A_1 & I_m \\ \vdots & \vdots \\ A_n & I_m \end{pmatrix}$$

$$Q = \sigma_a^2 (I_n \otimes E_m) + \sigma_b^2 (E_n \otimes I_m) + \sigma_e^2 I_{nm}$$

Äquivalente Formulierung:

$$y_{ij} = \mu + \alpha_i + \beta_j + a_i + b_j + e_{ij}, \quad i=1, \dots, n, \quad j=1, \dots, m;$$

$$y = X_1 \beta + X_2 b + e \text{ mit } \beta = (\mu, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)'$$

$$b = (a_1, \dots, a_n, b_1, \dots, b_m)'$$

$$e = (e_{11}, \dots, e_{1m}, \dots, e_{n1}, \dots, e_{nm})'$$

X_1, X_2 wie oben

Es ist $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$ und $\text{Rg } X_1 = \text{Rg } X_2 = n + m - 1$, also $\mathcal{M}(X_2) = \mathcal{M}(X_1)$.

P_L ist GMO zum Modell $M(L, Q)$ (siehe Abschnitt 1.3.0.)

Es ist $P_L = X_1 (X_1' X_1)^{-1} X_1' = \frac{1}{n} (I_n \otimes E_m) + \frac{1}{m} (E_n \otimes I_m) - \frac{1}{nm} E_{nm}$,

bzw. $P_L Y = (\frac{1}{m} Y_{i.} + \frac{1}{n} Y_{.j} - \frac{1}{nm} Y_{..})_{i=1, \dots, n, j=1, \dots, m}$

(Dabei bedeute "." Summation über alle Beobachtungen,

z.B. $Y_{.j} := \sum_i Y_{ij}$, $Y_{..} := \sum_{i,j} Y_{ij}$, etc.)

(Man sieht auch unmittelbar, daß $P_L Q = Q P_L$ symmetrisch ist, daß also P_L GMO zum Modell $M(L, Q)$ ist.)

Missing-Data-Problem

Annahme: Es fehle eine Beobachtung; oBdA sei dies y_{nm} ,
d.h. $P = \text{diag} (I_{nm-1}, 0)$.

Falls $n > 1$ und $m > 1$, so ist $\text{Rg}(P X_1) = \text{Rg}(X_1) = n + m - 1$, also $\dim P L = \dim L$. Da $\mathcal{M}(X_1) = \mathcal{M}(X_2)$ ist $\tilde{G} = P P_L = P P_L (I + A) P$ GMO zum Modell $M(P L, P Q P)$ (siehe Abschnitt 1.3.0.).

- 1) Direkte Berechnung von $\tilde{G} = P X_1 (X_1' (P Q P)^+ X_1)^{-1} X_1' (P Q P)^+$
bzw. $P_{PL} = P X_1 (X_1' P X_1)^{-1} X_1' P$:

Es ist

$$X_1' P X_1 = \begin{pmatrix} nm-1 & m & \dots & m & m-1 & n & \dots & n & n-1 \\ m & m & & & & 1 & \dots & 1 \\ \cdot & \cdot & & 0 & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & & & \cdot & \cdot & \cdot \\ \cdot & 0 & & \cdot & & \cdot & \cdot & \cdot \\ m & & & m & & \cdot & 1 & 1 \\ m-1 & & & & m-1 & 1 & \dots & 1 & 0 \\ n & 1 & \dots & 1 & & n & & & \\ \cdot & \cdot & \cdot & & \cdot & \cdot & & 0 & \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \\ n & \cdot & & 1 & 1 & 0 & & n & \\ n-1 & 1 & \dots & 1 & 0 & & & & n-1 \end{pmatrix}$$

Es ist $Rg(X_1' P X_1) = \dim P L = \dim L = n + m - 1$.

Zur direkten Berechnung von $\tilde{G}$ ist die Inverse einer Matrix vom Rang $\dim L$ neu zu berechnen!

(siehe auch anschließendes Zahlenbeispiel)

2) Berechnung von $P P_L (I + A) P$:

$$\text{Es ist } [(I-P)(I-P_L)(I-P)] = \frac{(n-1)(m-1)}{nm} (I-P)$$

$$\text{und } [(I-P)(I-P_L)(I-P)]^+ = \frac{nm}{(n-1)(m-1)} (I-P).$$

$$AP = [(I-P)(I-P_L)(I-P)]^+ P_L P = \frac{1}{(n-1)(m-1)} \cdot$$

$$\cdot \left(\begin{array}{c|c|c|c} 0 & \dots & 0 & 0 \\ \hline -1, \dots, -1, m-1 & \dots & -1, \dots, -1, m-1 & n-1, \dots, n-1, 0 \end{array} \right),$$

$\forall y \in V$ ist $x_o := APy = (0, \dots, 0, x_{nm})'$ mit

$$x_{nm} = \frac{1}{(n-1)(m-1)} \{ n y_{om} + m y_{no} - y_{oo} \}.$$

(Dabei bedeute "o" Summation über alle vorhandenen Beobachtungen:

$$y_{om} := \sum_{i=1}^{n-1} y_{im}, \quad y_{oo} := \sum_{j=1}^{m-1} y_{.j} + y_{om}, \text{ etc.})$$

(vergleiche KRUSKAL 1961)

Zahlenbeispiel: $n=2, m=3, \sigma_a^2=\sigma_b^2=\sigma_e^2=1$

$$X_1 = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \quad X_2 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

$$\left(\begin{array}{ccc|ccc} 3 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 3 & 1 & 1 \\ 0 & 1 & 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array} \right) = X_1 X_1' + I = Q$$

$$P_L = X_1 (X_1' X_1)^{-1} X_1' = \frac{1}{3} (I_2 \otimes E_3) + \frac{1}{2} (E_2 \otimes I_3) - \frac{1}{6} E_6 =$$

$$= \frac{1}{6} \cdot \left(\begin{array}{ccc|ccc} 4 & 1 & 1 & 2 & -1 & -1 \\ 1 & 4 & 1 & -1 & 2 & -1 \\ 1 & 1 & 4 & -1 & -1 & 2 \\ \hline 2 & -1 & -1 & 4 & 1 & 1 \\ -1 & 2 & -1 & 1 & 4 & 1 \\ -1 & -1 & 2 & 1 & 1 & 4 \end{array} \right)$$

Missing-Data-Problem: Sei $P = \text{diag}(I_{nm-1}, 0)$ wie oben.

1) Direkte Berechnung von $\tilde{G} = P X_1 [X_1' (P Q P)^+ X_1]^{-1} X_1' (P Q P)^+$

$$P Q P = \left(\begin{array}{ccc|ccc} 3 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 3 & 1 & 0 \\ 0 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \text{Rg}(P Q P) = 5$$

$$(PQP)^+ = \frac{1}{111} \cdot \left(\begin{array}{ccc|ccc} 55 & -19 & -12 & -23 & 14 & 0 \\ -19 & 55 & -12 & 14 & -23 & 0 \\ -12 & -12 & 45 & 3 & 3 & 0 \\ \hline -23 & 14 & 3 & 52 & -22 & 0 \\ 14 & -23 & 3 & -22 & 52 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$X_1'(PQP)^+X_1 = \frac{1}{111} \cdot \left(\begin{array}{ccc|ccc} 105 & 57 & 48 & 39 & 39 & 27 \\ \hline 57 & 69 & -12 & 18 & 18 & 21 \\ 48 & -12 & 60 & 21 & 21 & 6 \\ \hline 39 & 18 & 21 & 61 & -13 & -9 \\ 39 & 18 & 21 & -13 & 61 & -9 \\ 27 & 21 & 6 & -9 & -9 & 45 \end{array} \right)$$

$\text{Rg}(X_1'(PQP)^+X_1) = 4$, eine reguläre 4×4 -Untermatrix ist z.B.

$$W := \frac{1}{111} \begin{pmatrix} 105 & 57 & 39 & 39 \\ 57 & 69 & 18 & 18 \\ 39 & 18 & 61 & -13 \\ 39 & 18 & -13 & 61 \end{pmatrix} \quad W^{-1} = \frac{1}{4} \begin{pmatrix} 16 & -8 & -10 & -10 \\ -8 & 12 & 2 & 2 \\ -10 & 2 & 15 & 9 \\ -10 & 2 & 9 & 15 \end{pmatrix}$$

Eine mögliche Wahl für $\bar{X}_1'(PQP)^+X_1^-$ ist

$$\frac{1}{4} \cdot \left(\begin{array}{ccc|ccc} 16 & -8 & 0 & -10 & -10 & 0 \\ \hline -8 & 12 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -10 & 2 & 0 & 15 & 9 & 0 \\ -10 & 2 & 0 & 9 & 15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Damit ergibt sich für $\tilde{G} = PX_1 [X_1' (PQP)^+ X_1]^{-1} X_1' (PQP)^+$:

$$\tilde{G} = \frac{1}{4} \cdot \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & -1 & 0 \\ 1 & 3 & 0 & -1 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 3 & 1 & 0 \\ -1 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) .$$

Da $\tilde{G}$ symmetrisch ist, muß wegen Lemma 0(4) $\tilde{G} = P_{PL}$ sein.

2) Berechnung von $PP_L(I+A)P = P_{PL}$:

$$A = [(I-P)(I-P_L)(I-P)]^+ (I-P)P_L P =$$

$$= \frac{1}{2} \cdot \left(\begin{array}{ccc|ccc} & & & & & \\ & 0 & & & 0 & \\ \hline & -1 & -1 & 2 & 1 & 1 & 0 \end{array} \right) \quad (\text{siehe oben})$$

$$PP_L(I+A)P = \frac{1}{4} \cdot \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & -1 & 0 \\ 1 & 3 & 0 & -1 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 3 & 1 & 0 \\ -1 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) .$$

1.3.2. $\mathcal{M}(X_1) \subsetneq \mathcal{M}(X_2)$: Fehlende Daten sind nicht "durch Schätzwerte ersetzbar"

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

Sei $V := \mathbb{R}^{nm}$,
 $L := \mathcal{M}(X_1)$ mit $X_1 := \begin{pmatrix} 1 & A_1 \\ \vdots & \vdots \\ 1 & A_n \end{pmatrix}$, A_i wie in 1.3.1.

$Q := X_2 \Sigma X_2' + \sigma^2 I_{nm}$, X_2 und Σ wie in 1.3.1.

$Q = \sigma_a^2 (I_n \otimes E_m) + \sigma_b^2 (E_n \otimes I_m) + \sigma^2 I_{nm}$.

Äquivalente Formulierung:

$$y_{ij} = \mu + \alpha_i + a_i + b_j + e_{ij}, \quad i=1, \dots, n, \quad j=1, \dots, m;$$

(vergleiche Beispiel 1.3.1.)

Es ist $\mathcal{M}(X_1) \subsetneq \mathcal{M}(X_2)$ und $\text{Rg}(X_1) = n < n + m - 1 = \text{Rg}(X_2)$ - falls $m > 1$ - also $\mathcal{M}(X_1) \neq \mathcal{M}(X_2)$.

$P_L = X_1 (X_1' X_1)^{-1} X_1' = \frac{1}{m} (I_n \otimes E_m)$ ist GMO zum Modell $M(L, Q)$, da $P_L Q = Q P_L$ symmetrisch.

Missing-Data-Problem

Annahme: Es fehle eine Beobachtung, $P = \text{diag}(I_{nm-1}, 0)$
 wie in 1.3.1.;

dann ist $\text{Rg}(P X_1) = \text{Rg}(X_1) = n$, also $\dim PL = \dim L$.

Es ist $PP_L(I + A)P = P_{PL}$ nicht GMO zum Modell $M(PL, PQP)$.

Ich zeige: $P_{PL}PQP$ ist nicht symmetrisch.

$$\text{Es ist } P_{PL} = \text{diag} \left(\frac{1}{m}(I_{n-1} \otimes E_m), \frac{1}{m-1} E_{m-1}, 0 \right),$$

$$\begin{aligned} \text{und } P_{PL}P(I_n \otimes E_m)P &= \text{diag} \left((I_{n-1} \otimes E_m), E_{m-1}, 0 \right) = \\ &= P(I_n \otimes E_m)P \quad \text{symmetrisch,} \end{aligned}$$

$$P_{PL}P(E_n \otimes I_m)P \quad \text{jedoch nicht symmetrisch.}$$

Also ist für $\sigma_b^2 \neq 0$ und $Q = \sigma_a^2(I_n \otimes E_m) + \sigma_b^2(E_n \otimes I_m) + \sigma_e^2 I$
 $P_{PL}PQP$ nicht symmetrisch.

Wegen Lemma 0(4) ist damit P_{PL} nicht GMO zum Modell $M(PL, PQP)$.

Zahlenbeispiel: $n=2, m=3, \sigma_e^2 = \sigma_a^2 = \sigma_b^2 = 1$

$$X_1 = \left(\begin{array}{c|ccc} 1 & 1 & 0 & \\ 1 & 1 & 0 & \\ 1 & 1 & 0 & \\ \hline 1 & 0 & 1 & \\ 1 & 0 & 1 & \\ 1 & 0 & 1 & \end{array} \right), \quad X_2 \text{ und } Q \text{ wie in 1.3.1.}$$

$$P_L = \frac{1}{3} (I_2 \otimes E_3) = \frac{1}{3} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & & & \\ 1 & 1 & 1 & & 0 & \\ 1 & 1 & 1 & & & \\ \hline & & & 1 & 1 & 1 \\ & & & 1 & 1 & 1 \\ & 0 & & 1 & 1 & 1 \end{array} \right)$$

Berechnung von $\tilde{G} = PX_1[X_1'(PQP)^+X_1]^{-1}X_1'(PQP)^+$

$(PQP)^+$ wie in 1.3.1.,

$$X_1'(PQP)^+X_1 = \frac{1}{111} \begin{pmatrix} 105 & 57 & 48 \\ 57 & 69 & -12 \\ 48 & -12 & 60 \end{pmatrix}$$

$$\text{Rg} (X_1'(PQP)^+X_1) = 2.$$

Eine reguläre 2x2-Untermatrix ist

$$W := \frac{1}{111} \begin{pmatrix} 105 & 57 \\ 57 & 69 \end{pmatrix}, \quad W^{-1} = \frac{1}{36} \begin{pmatrix} 69 & -57 \\ -57 & 105 \end{pmatrix}.$$

Eine mögliche Wahl für $[X_1'(PQP)^+X_1]^{-}$ ist

$$\frac{1}{111} \begin{pmatrix} 69 & -57 & 0 \\ -57 & 105 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{G} = PX_1[X_1'(PQP)^+X_1]^{-}X_1'(PQP)^+ = \frac{1}{12} \begin{pmatrix} 4 & 4 & 4 & | & 0 \\ 4 & 4 & 4 & | & 0 \\ 4 & 4 & 4 & | & 0 \\ \hline -1 & -1 & 2 & | & 6 & 6 & 0 \\ -1 & -1 & 2 & | & 6 & 6 & 0 \\ 0 & 0 & 0 & | & 0 & 0 & 0 \end{pmatrix}$$

$\tilde{G}$ ist nicht symmetrisch, also ist $\tilde{G} \neq P_{PL}$ und

$PP_L(I+A)P = P_{PL}$ ist nicht GMO zum Modell $M(PL, PQP)$.

Es ist $P_{PL} = \text{diag}(\frac{1}{3} E_3, \frac{1}{2} E_2, 0) =$

$$= \frac{1}{6} \begin{pmatrix} 2 & 2 & 2 & | & 0 \\ 2 & 2 & 2 & | & 0 \\ 2 & 2 & 2 & | & 0 \\ \hline 0 & 0 & 0 & | & 3 & 3 & 0 \\ 0 & 0 & 0 & | & 3 & 3 & 0 \\ 0 & 0 & 0 & | & 0 & 0 & 0 \end{pmatrix}$$

1.3.3. $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$: Fehlende Daten sind "durch Schätzwerte ersetzbar"

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

Sei $V := \mathbb{R}^{nm}$,

$L := \mathcal{M}(X_1)$ mit X_1 wie in 1.3.1.,

$Q := X_2 \Sigma X_2' + \sigma_e^2 I_{nm}$ mit $X_2' = (A_1', \dots, A_n')$, A_i wie in 1.3.1.

und $\Sigma = \sigma_a^2 (I_n \otimes E_m)$,

$Q = \sigma_a^2 (I_n \otimes E_m) + \sigma_e^2 I_{nm}$.

Äquivalente Formulierung:

$$y_{ij} = \mu + \alpha_i + \beta_j + a_i + e_{ij}, \quad i=1, \dots, n, \quad j=1, \dots, m;$$

(vergleiche Beispiel 1.3.1.)

Wegen $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$ ist $P_L = \frac{1}{n}(I_n \otimes E_m) + \frac{1}{m}(E_n \otimes I_m) - \frac{1}{nm} E_{nm}$

GMO zum Modell $M(L, Q)$ und

$P_{PL} = PP_L(I + A)P$ GMO zum Modell $M(PL, PQP)$ für alle $P = P' = P^2$
mit $\dim PL = \dim L$.

Zahlenbeispiel: $n=2, m=3, \sigma_a^2=\sigma_e^2=1$

$$Q = I_6 + (I_2 \otimes E_3) = \left(\begin{array}{ccc|ccc} 2 & 1 & 1 & & & \\ 1 & 2 & 1 & & 0 & \\ 1 & 1 & 2 & & & \\ \hline & & & 2 & 1 & 1 \\ & 0 & & 1 & 2 & 1 \\ & & & 1 & 1 & 2 \end{array} \right)$$

P_L wie in 1.3.1.

Sei $P = \text{diag}(I_5, 0)$, so ist A und $PP_L(I+A)P = P_{PL}$ wie in 1.3.1.

Es ist

$$P_{PL}PQP = \frac{1}{4} : \left(\begin{array}{ccc|ccc} 7 & 5 & 4 & 1 & -1 & 0 \\ 5 & 7 & 4 & -1 & 1 & 0 \\ 4 & 4 & 8 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 7 & 5 & 0 \\ -1 & 1 & 0 & 5 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

symmetrisch, also muß $P_{PL} = \hat{G}$ sein.

Direkte Berechnung von $\hat{G} = PX_1[X_1'(PQP)^+X_1]^{-1}X_1'(PQP)^+$

$$(PQP)^+ = \frac{1}{12} \left(\begin{array}{ccc|ccc} 9 & -3 & -3 & & & \\ -3 & 9 & -3 & & 0 & \\ -3 & -3 & 9 & & & \\ \hline & & & 8 & -4 & 0 \\ & 0 & & -4 & 8 & 0 \\ & & & 0 & 0 & 0 \end{array} \right)$$

$$X_1' (PQP)^+ X_1 = \frac{1}{12} \left(\begin{array}{ccc|ccc} 17 & 9 & 8 & 7 & 7 & 3 \\ \hline 9 & 9 & 0 & 3 & 3 & 3 \\ 8 & 0 & 8 & 4 & 4 & 0 \\ \hline 7 & 3 & 4 & 17 & -7 & -3 \\ 7 & 3 & 4 & -7 & 17 & -3 \\ 3 & 3 & 0 & -3 & -3 & 9 \end{array} \right)$$

$\text{Rg} (X_1' (PQP)^+ X_1) = 4$, eine reguläre 4×4 -Untermatrix ist

$$W := \frac{1}{12} \left(\begin{array}{cccc} 17 & 9 & 7 & 7 \\ 9 & 9 & 3 & 3 \\ 7 & 3 & 17 & -7 \\ 7 & 3 & -7 & 17 \end{array} \right), \quad W^{-1} = \frac{1}{4} \left(\begin{array}{cccc} 12 & -8 & -6 & -6 \\ -8 & 12 & 2 & 2 \\ -6 & 2 & 7 & 5 \\ -6 & 2 & 5 & 7 \end{array} \right).$$

Eine mögliche Wahl für $[X_1' (PQP)^+ X_1]^{-1}$ ist

$$\frac{1}{4} \cdot \left(\begin{array}{ccc|ccc} 12 & -8 & 0 & -6 & -6 & 0 \\ \hline -8 & 12 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -6 & 2 & 0 & 7 & 5 & 0 \\ -6 & 2 & 0 & 5 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$\tilde{G} = P X_1 [X_1' (PQP)^+ X_1]^{-1} X_1' (PQP)^+ = \frac{1}{4} \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & -1 & 0 \\ 1 & 3 & 0 & -1 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 3 & 1 & 0 \\ -1 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$\tilde{G} = P_{PL}$ (siehe Beispiel 1.3.1.).

2. Gauß-Markov-Problem für das Modell $M(L, \theta)$

2.0. Existenz und Eindeutigkeit des GM-Schätzers

Wegen Corollar 2 zu Lemma 0(3) existiert ein GMO zum Modell $M(L, \theta)$ genau dann wenn $L \cap \sum QL^\perp = 0$. In diesem Fall ist G auf $L \oplus \sum QL^\perp = L + \sum QV$ eindeutig bestimmt.

In Übereinstimmung mit Kapitel 1 beschränke ich mich auf die Berechnung von Projektoren $G: V \rightarrow L$ entlang $W := \sum QL^\perp \oplus (L + \sum QV)^\perp = QL^\perp \oplus (L^\perp \cap K(Q))$. G hängt von dem gegebenen Skalarprodukt ab.

2.1. Berechnung eines GM-Schätzers

Lemma 2.1(1) Sei $L \cap \sum QL^\perp = 0$, $Q_0 \in \theta$ regulär

$$\implies M(L, \theta) \approx M(L, Q_0)$$

Beweis: $Q_0 \in \theta$ regulär, so gilt:

$$V = Q_0 V \subseteq \sum QV \subseteq V \quad \text{und} \quad Q_0 L^\perp \subseteq \sum QL^\perp.$$

Da $V = L \oplus Q_0 L^\perp \subseteq L \oplus \sum QL^\perp \subseteq V$, folgt

$$\sum QL^\perp = Q_0 L^\perp.$$

Corollar: Sei $Q_0 \in \theta$ regulär, G GMO zum Modell $M(L, Q_0)$, so sind folgende Aussagen paarweise äquivalent:

i) $L \cap \sum QL^\perp = 0$

ii) G ist GMO zum Modell $M(L, \theta)$

iii) $Q\bar{Q}$ symmetrisch $\forall Q \in \bar{\theta}$

iv) $M(L, \bar{\theta}) \cong M(L, Q_0)$

Offensichtlich gibt es für beliebiges $\theta \subseteq S_+(V)$ kein $\bar{Q} \in S_+(V)$ so daß $M(L, \theta) \cong M(L, \bar{Q})$.

Ich bestimme zunächst eine maximale Menge $\bar{\theta} \subseteq S_+(V)$ so daß $M(L, \bar{\theta}) \cong M(L, \theta)$.

Die Menge $S(V)$ aller selbstadjungierten linearen Abbildungen von V in sich mit den übliche Operationen ist ein endlich-dimensionaler reeller linearer Raum und $S_+(V) \subseteq S(V)$.

Lemma 2.1(2) Sei $\theta \subseteq S_+(V)$, $[\theta]$ die lineare Hülle von θ in $S(V)$

$\Rightarrow \exists Q_1, \dots, Q_m \in \theta$ so daß
 $\{Q_1, \dots, Q_m\} =: \theta_0$ Basis von $[\theta]$.

Im Folgenden sei stets $\theta_0 = \{Q_1, \dots, Q_m\} \subseteq \theta$ so daß $[\theta_0] = [\theta]$
 (m nicht notwendig minimal) und $\bar{\theta} := [\theta] \cap S_+(V)$.

Lemma 2.1(3) Für einen beliebigen linearen Teilraum $M \subseteq V$ gilt:

$$\sum_{i=1}^m Q_i M = \sum_{Q \in \bar{\theta}} QM = \sum_{Q \in \theta} QM = \sum_{Q \in [\theta]} QM.$$

Beweis: " $\subseteq$ " folgt aus $\theta_0 \subseteq \theta \subseteq \bar{\theta} \subseteq [\theta]$

" $\supseteq$ ": es reicht zu zeigen:

$$\sum_{Q \in \bar{\theta}} QM \subseteq \sum_{i=1}^m Q_i M.$$

Dieses folgt aber aus $[\theta_0] = [\theta]$.

Corollar 1 $M(L, \theta_0) \cong M(L, \theta) \cong M(L, \bar{\theta})$

Beweis: folgt aus Lemma 2.1(3) mit $M=V$
bzw. $M=L^\perp$.

Corollar 2 $\sum_{i=1}^m K(Q_i) = \sum_{Q \in \bar{\theta}} K(Q) = \sum_{Q \in \bar{\theta}} K(Q) = \sum_{Q \in [\bar{\theta}]} K(Q).$

Beweis: folgt aus Lemma 2.1(3) mit
 $(\sum QV)^\perp = \bigcap K(Q).$

Lemma 2.1(4) Sei $\bar{\theta} := \{ \bar{Q} \in \bar{\theta} \mid \bar{Q}V = \sum QV \}$, so gilt:

- i) $Q_0 \in \bar{\theta}$ regulär $\implies Q_0 \in \bar{\theta}$
- ii) $\{Q = \sum_{i=1}^m c_i Q_i \mid c_i > 0\} \subseteq \bar{\theta}$

Beweis: ii) Sei $\bar{Q} = \sum c_i Q_i$ mit $c_i > 0$
zu zeigen: $K(\bar{Q}) = \bigcap K(Q_i)$

Sei $x \in V$, $Q_i x = 0 \quad \forall i \implies \bar{Q}x = 0.$

Sei $\bar{Q}x = 0 \implies 0 = (\bar{Q}x, x) = \sum c_i (Q_i x, x)$
da $c_i > 0$, muß $(Q_i x, x) = 0 \quad \forall i$ sein,
also $Q_i x = 0 \quad \forall i.$

Corollar: Für beliebiges $\theta \in S_+(V)$, $\theta \neq \phi$ ist $\bar{\theta} \neq \phi.$

Theorem 2(1) Sei $\theta \in S_+(V)$, $\theta \neq \emptyset$

$$\bar{\theta} := \{\bar{Q} \in [\theta] \cap S_+(V) \mid \bar{Q}V = \sum QV\},$$

so sind folgende Aussagen paarweise äquivalent:

- i) $L \cap \sum QL^\perp = 0$
- ii) $\exists \bar{Q} \in \bar{\theta}: M(L, \theta) \cong M(L, \bar{Q})$
- iii) $\forall \bar{Q} \in \bar{\theta}: M(L, \theta) \cong M(L, \bar{Q})$

Beweis: es ist nur zu zeigen: i) $\Rightarrow$ iii)

Sei $\bar{Q} \in \bar{\theta}$ beliebig, so ist $\bar{Q}V = \sum QV$,

zu zeigen: $\bar{Q}L^\perp = \sum QL^\perp$.

$$\text{Wegen } \bar{Q} \in [\theta] \text{ ist } \bar{Q}L^\perp \subseteq \overline{\sum_{Q \in [\theta]} QL^\perp} = \sum_{Q \in \bar{\theta}} QL^\perp$$

(siehe Lemma 2.1(3)).

Es bleibt zu zeigen: $QL^\perp \subseteq \bar{Q}L^\perp \quad \forall Q \in \theta$.

Sei $Q \in \theta$, $a \in L^\perp$ beliebig mit $Qa \neq 0$.

Wegen $V = L \oplus \bar{Q}L^\perp \oplus (L^\perp \cap K(\bar{Q}))$

und $K(\bar{Q}) = \bigcap K(Q)$ nach Wahl von $\bar{Q}$

hat Qa eine Darstellung als

$$Qa = l + \bar{Q}b \text{ mit } l \in L \text{ und } b \in L^\perp.$$

$$\Rightarrow l = Qa - \bar{Q}b \in L \cap \sum QL^\perp = 0 \text{ nach Vor.},$$

$$\text{also } Qa = \bar{Q}b \in \bar{Q}L^\perp.$$

Corollar:

Sei $\bar{Q} \in \bar{\theta}$ beliebig, G GMO zum Modell $M(L, \bar{Q})$,

so sind folgende Aussagen paarweise äquivalent:

- i) $L \cap \sum QL^\perp = 0$
- ii) G ist GMO zum Modell $M(L, \theta)$
- iii) GQ symmetrisch $\forall Q \in \theta$
- iv) $M(L, \theta) \cong M(L, \bar{Q})$.

(vergleiche DRYGAS 1977, Corollar 2.2.)

Bemerkung: Falls eine der oben genannten Bedingungen erfüllt ist, hängt G nicht mehr von dem speziell gewählten $\bar{Q} \in \bar{\Theta}$ ab.

Anwendung: Falls ein reguläres $Q_0 \in \Theta$ bekannt ist oder $Q_1, \dots, Q_m \in \Theta$ so daß für $\Theta_0 := \{Q_1, \dots, Q_m\}$ gilt $[\Theta_0] = [\Theta]$ (m nicht notwendig minimal), so kann man mit $\bar{Q} := Q_0$ bzw. $\bar{Q} := \sum Q_i$ prüfen, ob das GM-Problem für das Modell $M(L, \Theta)$ eine Lösung besitzt:

Mit den Formeln aus Kapitel 1. berechnet man einen Projektor $G: V \rightarrow L$, GMO zum Modell $M(L, \bar{Q})$ und prüft, ob GQ für alle $Q \in \Theta$ symmetrisch ist (bzw. hinreichend: GQ_i symmetrisch für $i=1, \dots, m$).

Falls GQ für alle $Q \in \Theta$ symmetrisch ist, so ist G bereits eine mögliche Lösung. Falls GQ nicht symmetrisch für mindestens ein $Q \in \Theta$, so existiert kein GMO zum Modell $M(L, \Theta)$.

Statt $\bar{Q} \in \bar{\Theta}$ kann man entsprechend Theorem 1(5) bzw. 1(6) auch $T := \bar{Q} + AA'$ (falls $L = AU$ - siehe oben) oder $\tilde{Q} := T + (I - TT^+)$ betrachten und genauso verfahren.

Im Hinblick auf die Anwendung auf quadratische Modelle (siehe Kapitel 3.) sind folgende Kriterien nützlich:

Lemma 2.1(5) Für beliebiges $\bar{Q} \in \bar{\Theta}$ gilt:

$$L \cap \sum QL^\perp = 0 \quad \Longleftrightarrow \quad Q \bar{Q}^{-1} L \subseteq L \quad \forall Q \in \Theta$$

Beweis: Wegen Theorem 2(1) ist

$$L \cap \sum QL^\perp = 0 \quad \Longleftrightarrow \quad QL^\perp \subseteq \bar{Q}L^\perp \quad \forall Q \in \Theta$$

Mit Farkas' Theorem ist äquivalent:

$$\bar{Q}^{-1}L = (\bar{Q}L^{\perp})^{\perp} \subseteq (QL^{\perp})^{\perp} = Q^{-1}L,$$

(vergleiche auch DRYGAS 1977, Theorem 2.1.)

Bemerkung: Falls $I \in [0]$, so ergibt sich mit $\bar{Q} = I$:

$$L \cap \sum QL^{\perp} = 0 \iff QL \subseteq L \quad \forall Q \in \Theta$$

(siehe Corollar zu Lemma 0(4), vergleiche auch EATON 1978 Theorem 3.3.)

Corollar 1 Sei $\bar{Q}^{+}$ die eindeutig bestimmte Moore-Penrose-Inverse zu $\bar{Q}$, so gilt:

$$Q \bar{Q}^{+}L \subseteq L \quad \forall Q \in \Theta \implies L \cap \sum QL^{\perp} = 0.$$

Beweis: ich zeige: $Q \bar{Q}^{-1}L \subseteq L \quad \forall Q \in \Theta$.

$$\text{Sei } \bar{Q}x \in L \quad (x \in \bar{Q}^{-1}L)$$

$$\implies Qx = Q\bar{Q}^{+}\bar{Q}x \quad \text{nach Corollar 2 zu Lemma 2.1(3),}$$

$$\text{wegen } Q\bar{Q}^{+}L \subseteq L \quad \text{also } Qx \in L.$$

Corollar 2 Falls $L \subseteq \sum QV$, $\bar{Q} \in \bar{\Theta}$ beliebig (d.h. $\bar{Q}$ quasiregulär bezüglich L), so gilt:

$$L \cap \sum QL^{\perp} = 0 \implies Q\bar{Q}^{+}L \subseteq L \quad \forall Q \in \Theta$$

Beweis: $L \subseteq \bar{Q}V \implies \bar{Q} \bar{Q}^{+}L = L$, also $\bar{Q}^{+}L \subseteq \bar{Q}^{-1}L$
und $Q\bar{Q}^{+}L \subseteq Q\bar{Q}^{-1}L \quad \forall Q \in \Theta$.

$$\text{Aus } L \cap \sum QL^{\perp} = 0 \text{ folgt mit Lemma 2.1(5)}$$

$$Q\bar{Q}^{-1}L \subseteq L \quad \forall Q \in \Theta,$$

$$\text{also insgesamt } Q\bar{Q}^{+}L \subseteq L \quad \forall Q \in \Theta.$$

Damit ist für den Fall $L \subseteq \sum QV$ ein notwendiges und hinreichendes Kriterium für die Existenz gleichmäßig bester Schätzer gegeben.

2.2. Missing-Data-Problem - Das Modell $M(\mathbb{P}L, \mathbb{P}\odot\mathbb{P})$

Es sei $\mathbb{P} = \mathbb{P}' = \mathbb{P}^2 \in S_+(V)$ wie in Kapitel 1.2.0. und

$$\mathbb{P}\odot\mathbb{P} := \{\mathbb{P}Q\mathbb{P} \mid Q \in \Theta\}.$$

Im Falle fehlender Daten ist als neues Modell $M(\mathbb{P}L, \mathbb{P}\odot\mathbb{P})$ in $\mathbb{P}V$ bzw. in V zu betrachten (siehe Lemma 1.2(2)).

Offensichtlich ist die Bedingung $M(L, \Theta) \cong M(L, \bar{Q})$ ($\bar{Q} \in \bar{\Theta}$) nicht hinreichend dafür, daß $M(\mathbb{P}L, \mathbb{P}\odot\mathbb{P}) \cong M(\mathbb{P}L, \mathbb{P}\bar{Q}\mathbb{P})$; i.a. braucht für $M(\mathbb{P}L, \mathbb{P}\odot\mathbb{P})$ überhaupt kein GMO zu existieren (siehe Beispiel 1.3.2. bzw. Absatz 3) der Ergänzungen zu 1.2.).

Lemma 2.2(1) Für alle $\bar{Q} \in \bar{\Theta}$ gilt: $\mathbb{P}\bar{Q}\mathbb{P}V = \sum \mathbb{P}Q\mathbb{P}V$

$$\text{d.h. } \mathbb{P}\bar{\Theta}\mathbb{P} \subseteq \overline{\mathbb{P}\odot\mathbb{P}}.$$

$$\text{Beweis: } K(\mathbb{P}\bar{Q}\mathbb{P}) = [K(\bar{Q}) \cap \mathbb{P}V] \oplus V_2 \quad (\text{Lemma 1.2(1)})$$

$$= [\bigcap K(Q) \cap \mathbb{P}V] \oplus V_2 =$$

$$= \bigcap [(K(Q) \cap \mathbb{P}V) \oplus V_2] = \bigcap_{Q \in \Theta} K(\mathbb{P}Q\mathbb{P}).$$

Mit Farkas' Theorem folgt die Behauptung.

Anwendung: Mit demselben $\bar{Q}$ wie für das vollständige Modell $M(L, \Theta)$ ist also zu prüfen, ob auch für $M(\mathbb{P}L, \mathbb{P}\odot\mathbb{P})$ ein GMO existiert.

Falls $\dim \mathbb{P}L = \dim L$, so ist der GMO $\tilde{G}$ für das Modell $M(\mathbb{P}L, \mathbb{P}\bar{Q}\mathbb{P})$ eventuell mit Hilfe des für das vollständige Modell $M(L, \bar{Q})$ bekannten Projektors G zu berechnen (siehe Theorem 1(8)).

Theorem 2(2) Sei $\Theta \in S_+(V)$, $\bar{\Theta}$ wie in Theorem 2(1),
 $\bar{Q} \in \bar{\Theta}$ beliebig und $M(L, \Theta) \cong M(L, \bar{Q})$.
 Ferner sei $P = P' = P^2$ wie in 1.2.0.

Dann sind folgende Aussagen äquivalent:

- i) Es existiert ein GMO zum Modell $M(PL, P\Theta P)$
- ii) $M(PL, P\Theta P) \cong M(PL, P\bar{Q}P)$

Corollar:

Voraussetzungen wie in Theorem 2(2).

G sei GMO zum Modell $M(L, \bar{Q})$,

Für $M(L, \bar{Q})$ und P sei eine der Aussagen von
 Theorem 1(8) iv) erfüllt.

$\tilde{G} := PG(I+A)P$ mit

$$A := - [(I-P)(I-G)(I-P)]^+ (I-G)P,$$

so sind folgende Aussagen paarweise äquivalent:

- i) Es existiert ein GMO zum Modell $M(PL, P\Theta P)$
- ii) $\tilde{G}$ ist GMO zum Modell $M(PL, P\Theta P)$
- iii) $\tilde{G}P\Theta P$ symmetrisch $\forall Q \in \Theta$.

Anmerkungen zu 2.1. und 2.2.

- 1) Sehr allgemeine (notwendige und hinreichende) Bedingungen für die Existenz gleichmäßig bester Schätzer zum Modell $M(L, \Theta)$ wurden von Drygas (DRYGAS 1976, Theorem 1.2.) und Eaton (EATON 1978, Theorem 3.1.) angegeben.

Für spezielles Θ wurde dieses Problem zuvor von Kruskal (KRUSKAL 1968, Theorem 1 - $\Theta := \{I, Q\}$) sowie Mitra & Moore (MITRA & MOORE 1973 - $\Theta := \{Q, Q_0\}$) behandelt.

Unter etwas modifizierter Fragestellung erscheint eine Lemma 2.1(5) entsprechende Formulierung bei Rehder (REHDER 1980, Corollar zu Theorem 3.1.)

Alle angegebenen Bedingungen haben die Form von Beziehungen zwischen linearen Teilräumen von V . Für die praktische Anwendung sind sie kaum geeignet, zumal sie keinen Hinweis auf Berechnungsmöglichkeiten des optimalen Schätzers geben.

Drygas entwickelt in seiner 1977 erschienen Arbeit das Konzept eines "maximalen" $\bar{Q}$ ($\bar{Q}V \supseteq QV$). Er zeigt, daß im Falle der Existenz eines GMO $\bar{Q}L^\perp \supseteq QL^\perp$ - also $M(L, \Theta) \leq M(L, \bar{Q})$ - ist (siehe DRYGAS 1977, Corollar 2.2.a)).

Meine Darstellungen ermöglichen es, ein $\bar{Q}$ so anzugeben, daß $\bar{Q}V = \sum QV$ ist. Im Falle der Existenz eines GMO gilt dann $\bar{Q}L^\perp = \sum QL^\perp$ - also $M(L, \Theta) \cong M(L, \bar{Q})$. Damit ist der zu berechnende Projektor G im Sinne von Kapitel 1.1. eindeutig festgelegt (Theorem 2.(1)).

- 2) Die Frage nach der Existenz gleichmäßig bester Schätzer im Falle fehlender Daten stellt kein prinzipiell neues Problem dar.

Da man in der Praxis jedoch den Versuchsplan nach Möglichkeit so anlegt, daß für das zugehörige (vollständige) Modell $M(L, \theta)$ ein GMO existiert, also $M(L, \theta) \cong M(L, \bar{Q})$ ist, liegt es nahe zu untersuchen, welche Veränderungen das Modell beim Verlust einzelner Meßdaten erfährt und wie sich diese Veränderungen in Bezug auf das GM-Problem auswirken.

Die Kapitel 1.1. - 2.2. geben eine anschaulich geometrische Beschreibung:

Durch $\bar{Q}$ wird eine Norm $\|\dots\|_0$ auf V induziert (Theorem 1(6) und Absatz 4) der Anmerkungen zu 1.2.).

Falls $M(L, \theta) \cong M(L, \bar{Q})$, so ist $\sum Q L^\perp \subseteq L^{\perp_0}$ (Theorem 2(1)).

Der Verlust von Beobachtungen bedeutet geometrisch:

- i) eine Drehung des linearen Teilraumes $L \rightarrow PL$ in V und
- ii) Übergang zu der durch $P\bar{Q}P$ induzierten Norm $\|\dots\|_{00}$ auf V (siehe Theorem 1(8) und Theorem 2(2)).

Nur falls $\sum P Q P (PL)^\perp \subseteq (PL)^{\perp_{00}}$, existiert ein GMO zu $M(PL, P\bar{Q}P)$ (Theorem 2(2)).

2.3. Beispiele

2.3.0. Regressionsmodell mit festen und zufälligen Effekten

Sei $V := \mathbb{R}^n$, $L := \mathcal{H}(X_1)$, wobei X_1 eine bekannte reelle $n \times p$ -Matrix sei,

$\Theta := \{ Q = X_2 \Sigma X_2' + \sigma^2 I \mid \Sigma = \Sigma' \geq 0 \text{ } q \times q\text{-Matrix, } \sigma^2 \geq 0 \}$,
 X_2 bekannte reelle $n \times q$ -Matrix
 (vergleiche Beispiel 1.3.0. und 2.3.0.)

Behauptung: $\mathcal{H}(X_2) \subseteq \mathcal{H}(X_1)$, $P = P' = P^2$ beliebig, so gilt:

- i) $P_L = X_1 (X_1' X_1)^{-1} X_1'$ ist GMO zum Modell $M(L, \Theta)$
- ii) $P_{PL} = P X_1 (X_1' P X_1)^{-1} X_1' P$ ist GMO zum Modell $M(PL, P \Theta P)$.

Beweis: siehe Beispiel 1.3.0.

Insbesondere existiert auch zum reduzierten Modell $M(PL, P \Theta P)$ ein GMO.

Falls $\text{Rg}(P X_1) = \text{Rg}(X_1)$, so ist zusätzlich

$P_{PL} = P P_L (I - [(I-P)(I-P_L)(I-P)]^+ (I-P_L))$, d.h. fehlende Daten sind "durch Schätzwerte zu ersetzen".

2.3.1. $\mathcal{M}(X_1) = \mathcal{M}(X_2)$: GM-Problem für $M(\mathbf{P}_L, \mathbf{P} \otimes \mathbf{P})$ lösbar
durch $\tilde{G} = \mathbf{P}G(\mathbf{I} + \mathbf{A})\mathbf{P}$

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

Sei $V := \mathbb{R}^{nm}$, $L := (X_1)$

$$\Theta := \{Q = \sigma_a^2 (\mathbf{I}_n \otimes \mathbf{E}_m) + \sigma_b^2 (\mathbf{E}_n \otimes \mathbf{I}_m) + \sigma_e^2 \mathbf{I}_{nm} \mid \sigma^2 \geq 0\}$$

X_1, X_2 wie in 1.3.1.

Es ist $P_L Y = (\frac{1}{m} Y_{i.} + \frac{1}{n} Y_{.j} - \frac{1}{nm} Y_{..})_{i=1, \dots, n; j=1, \dots, m}$;

gleichmäßig bester Schätzer für $E_P Y$ für alle $\sigma_a^2, \sigma_b^2, \sigma_e^2 \geq 0$.

Sei $\mathbf{P} = \text{diag}(\mathbf{I}_{nm-1}, 0)$ wie in 1.3.1., so ist

$$\mathbf{P}_{PL} \mathbf{P} Y = \mathbf{P} \mathbf{P}_L (\mathbf{P} Y + x_0) \text{ mit } x_0 = (0, \dots, 0, \frac{1}{(n-1)(m-1)} (n y_{om} + m y_{no} - y_{oo}))'$$

gleichmäßig bester Schätzer für $E_P \mathbf{P} Y$ für alle $\sigma_a^2, \sigma_b^2, \sigma_e^2 \geq 0$.

2.3.2. $\mathcal{M}(X_1) \subsetneq \mathcal{M}(X_2)$: GM-Problem für $M(\mathbf{P}_L, \mathbf{P} \otimes \mathbf{P})$ nicht
lösbar

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

Sei $V := \mathbb{R}^{nm}$, $L := \mathcal{M}(X_1)$ mit X_1 wie in 1.3.2. und Θ wie in 2.3.1.

$$\text{Es ist } P_L Y = \frac{1}{m} (\mathbf{I}_n \otimes \mathbf{E}_m) Y = (\frac{1}{m} Y_{i.})_{i=1, \dots, n; j=1, \dots, m};$$

gleichmäßig bester Schätzer für $E_P Y$.

Sei $P = \text{diag}(I_{nm-1}, 0)$ wie in 1.3.2..

Wegen $\bar{Q} = I \in \Theta$ existiert ein GMO zum Modell $M(PL, P\Theta P)$ genau dann wenn $M(PL, P\Theta P) \cong M(PL, P)$ (Lemma 2.2(1) und Theorem 2(1)) bzw. $P_{PL} P\Theta P$ symmetrisch für alle $Q \in \Theta$ (Corollar 1 zu Theorem 2(1)).

In 1.3.2. wurde jedoch gezeigt, daß für $\sigma_b^2 \neq 0$ $P_{PL} P\Theta P$ nicht symmetrisch ist.

2.3.3. $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$: GM-Problem für $M(PL, P\Theta P)$ lösbar durch $\tilde{G} = PG(I + A)P$

Beispiel: Zweifache Kreuzklassifikation ohne Wechselwirkung

Sei $V := \mathbb{R}^{nm}$, $L := \mathcal{M}(X_1)$, X_1 wie in 1.3.3.

$$\Theta := \{ Q = \sigma_a^2 (I_n \otimes E_m) + \sigma_e^2 I_{nm} \mid \sigma_a^2, \sigma_e^2 \geq 0 \}$$

Es ist $P_L Y$ gleichmäßig bester Schätzer für $E_P Y$ für alle $\sigma_a^2, \sigma_e^2 \geq 0$ (P_L wie in 2.3.1.).

Sei $P = \text{diag}(I_{nm-1}, 0)$ wie in 1.3.3., so ist

$P_{PL} P Y = P P_L (P Y + x_0)$ gleichmäßig bester Schätzer für $E_P P Y$ für alle $\sigma_a^2, \sigma_e^2 \geq 0$; dabei ist x_0 wie in 2.3.1.

3. Gauß-Markov-Schätzung in quadratischen Modellen

3.0. Zurückführung auf ein lineares Problem

Sei $y: \Omega \rightarrow V$ ein Zufallsvektor wie in Kapitel 0. mit

$\text{Cov}_P Y = \sum_{i=1}^m \sigma_i^2 \cdot W_i$, wobei $W_1, \dots, W_m \in S_+(V)$ bekannte lineare

Operatoren seien.

$\sigma_1^2, \dots, \sigma_m^2$ sollen durch eine quadratische Funktion in y erwartungstreu "gleichmäßig optimal" geschätzt werden.

Setze $Y := yy'$, wobei $y = (y_1, \dots, y_n)'$ und $y_i = (y, e_i) \in \mathbb{R}$ die Koordinaten von y bei gegebener fester Basis von V seien.

Dann ist $E_P Y = \mu\mu' + \sum \sigma_i^2 \cdot W_i =: W_0$ ($\mu = E_P y$).

Unter der Annahme einer Quasi-Normalverteilung für y ist $\text{Cov}_P Y = 2(W_0 \otimes W_0 - \mu\mu' \otimes \mu\mu')$.

Neben dem Modell $M(L, \Theta)$ mit $[\Theta] = [W_1, \dots, W_m]$ zum Schätzen von $E_P y$ ist das Modell $M(L^*, \Theta^*)$ - lineares Modell in Y - in $S(V)$ zu betrachten mit

$L^* := \{ \mu\mu' + \sum c_i W_i \mid \mu \in L, c_i \in \mathbb{R} \}$ und

$\Theta^* := \{ W \otimes W - A \otimes A \mid W \in L^*, A = \mu\mu' \text{ mit } \mu \in L \}$

3.1. Gauß-Markov-Problem für das Modell $M(L^*, \Theta^*)$

3.1.0. Koordinatenfreie Formulierung

Sei V - wie in Kapitel 0. - ein reeller linearer Raum mit vorgegebenem Skalarprodukt $(\cdot, \cdot): V \rightarrow \mathbb{R}$ und $\dim V = n$.

Für $x, z \in V$ bezeichne $x \otimes z$ das (abstrakte) Tensorprodukt von x mit z und $V \otimes V := [\{x \otimes z \mid x, z \in V\}]$ den zugehörigen reellen Vektorraum (siehe HALMOS §24,25).

Wegen $\dim V \otimes V = n^2$ - falls $\{e_1, \dots, e_n\}$ Basis von V , so ist $\{e_i \otimes e_j \mid i, j = 1, \dots, n\}$ Basis von $V \otimes V$ - gibt es natürliche Isomorphismen:

$$i_1: V \otimes V \cong \mathcal{L}(V, V) - x \otimes z \mapsto A \in \mathcal{L}(V, V) \text{ mit } Av := (v, z)x,$$

(bei fest gewählter Basis von V .)

$$i_2: V \otimes V \cong \mathbb{R}^{n^2} - x \otimes z = (x_1, \dots, x_n)' \otimes z \mapsto ((x_1 z)', \dots, (x_n z)')' \\ \text{(Kronecker-Produkt der Vektoren } x \text{ und } z)$$

$$i_3: V \otimes V \cong \mathbb{R}_{n,n} - x \otimes z = \sum x_i z_j (e_i \otimes e_j) \mapsto (x_i z_j)_{i,j} = xz'.$$

$x \otimes z$ ist also - je nach Zusammenhang bzw. Anwendung - als Vektor von $V \otimes V$, als lineare Abbildung von V in sich, als reelles n^2 -Tupel oder als reelle $n \times n$ -Matrix zu interpretieren.

Allgemein seien die Elemente von $V \otimes V$ mit $A, B, \dots$ bezeichnet.

Das Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V läßt sich durch die Festlegung $\langle (x_1 \otimes z_1), (x_2 \otimes z_2) \rangle := (x_1, x_2) (z_1, z_2)$ in naheliegender Weise zu einem Skalarprodukt auf $V \otimes V$ erweitern:

Falls $\{e_1, \dots, e_n\}$ Basis von V mit $(e_i, e_j) = \delta_{ij}$ so daß für $x = \sum x_i e_i$ und $z = \sum z_i e_i \in V$ $(x, z) = \sum x_i z_i$ gilt,

so ist für $u = \sum u_{ij} (e_i \otimes e_j)$ und $v = \sum v_{ij} (e_i \otimes e_j) \in V \otimes V \cong \mathbb{R}^{n^2}$
 $\langle u, v \rangle = \sum u_{ij} v_{ij}$.

Für beliebige $A, B \in \mathcal{L}(V, V)$ (bzw. $\mathbb{R}_{n,n}$), $x, z \in V$ gilt dann

$$\langle A, B \rangle = \text{tr } AB' \text{ sowie } \langle A, (x \otimes z) \rangle = (Ax, z).$$

(siehe auch PUCKELSHEIM 1977)

In $V \otimes V$ sei ferner ein äußeres Produkt definiert durch

$$(A \otimes B)(x \otimes z) := Ax \otimes Bz \quad (x, z \in V).$$

Für beliebige $A, B, C \in \mathcal{L}(V, V)$ ist dann $(A \otimes B)C = ACB'$.

Für $A, B \in \mathbb{R}_{n,n}$ entspricht $A \otimes B$ dem Kronecker-Produkt der Matrizen.

3.1.1. Das Modell $M(L^*, \mathbb{C}^*)$

Lemma 3.1(1) Für $x \in V$ ist $x \otimes x \in S_+(V)$ und $\text{Rg } xx' \leq 1$

Beweis: siehe Halmos §51, Theorem 1

Corollar: $[\{ x \otimes x \mid x \in V \}] = S(V)$

Beweis: Wegen Lemma 3.1(1) ist

$$[\{ x \otimes x \mid x \in V \}] \subseteq [S_+(V)] \subseteq S(V).$$

Für jedes $A \in S(V)$ gilt:

$$A = \sum \alpha_i (a_i \otimes a_i) \quad \text{mit } a_i \in V \text{ Eigenvektor} \\ \text{zum Eigenwert } \alpha_i \in \mathbb{R}.$$

Sei $y: \Omega \rightarrow V$ eine Abbildung so daß für alle $x \in V$ (x, y) eine reelle Zufallsvariable ist.

$Y: \Omega \rightarrow V^* := S(V)$ sei definiert durch $Y(\omega) := y(\omega) \otimes y(\omega) \quad \forall \omega \in \Omega$,
dann ist für beliebiges $A \in V^* \quad \langle A, Y \rangle := \text{tr } AY = (A y, y): \Omega \rightarrow \mathbb{R}$
eine reelle Zufallsvariable.

Lemma 3.1(2) Es sei $\mathcal{P}$ eine Familie von Wahrscheinlichkeitsmaßen, so gilt:

- i) $\forall P \in \mathcal{P} \quad \exists$ genau ein $W \in V^*$: $E_P(\text{tr } AY) = \text{tr } AW$
für alle $A \in V^*$
- ii) $\forall P \in \mathcal{P} \quad \exists$ genau ein $Q^* \in S_+(V^*)$:
 $\text{Cov}_P((A y, y), (B y, y)) = \text{tr } A Q^* B \quad \forall A, B \in V^*.$

Beweis: siehe Lemma 0(1)

Bezeichnung: $W := E_P Y$ und $Q^* := \text{Cov}_P Y$

Falls $\mathcal{P}$ eine Familie von Quasi-Normalverteilungen ist,

$$L := [\{ E_{\mathcal{P}} Y \mid P \in \mathcal{P} \}] ,$$

$$\Theta := \{ Q = \text{Cov}_{\mathcal{P}} Y \mid P \in \mathcal{P} \} = \{ \sum \sigma_i^2 \cdot W_i \mid \sigma_i^2 \geq 0 \} \quad \text{mit } W_i = W_i' \geq 0,$$

so ist

$$(i=1, \dots, m)$$

$$L^* := [\{ E_{\mathcal{P}} Y \mid P \in \mathcal{P} \}] = [\{ 1 \otimes 1 \mid 1 \in L \}] + [W_1, \dots, W_m]$$

$$=: L_1 + L_2 \quad \text{mit } L_1 = S(L)$$

$$\Theta^* := \{ Q^* = \text{Cov}_{\mathcal{P}} Y \mid P \in \mathcal{P} \} \subseteq \{ 2(W \otimes W - A \otimes A) \mid W \in L^*, A \in L_1 \}$$

$$\subseteq [\{ W \otimes W \mid W \in L^* \}] \subseteq [\Theta^*]$$

Wegen Corollar 1 zu Lemma 2.1(3) ist für beliebiges $\Theta^* \subseteq S_+(V^*)$

$$M(L^*, \Theta^*) \cong M(L^*, \bar{\Theta}^*) \quad \text{mit } \bar{\Theta}^* := [\Theta^*] \cap S_+(V^*) .$$

Im Folgenden werde stets das Modell $M(L^*, \Theta^*)$ in V^* betrachtet mit

$$V^* := S(V) \subseteq V \otimes V,$$

$$L^* := L_1 + L_2 = S(L) + [W_1, \dots, W_m]$$

$$\Theta^* := [\Theta^*] \cap S_+(V^*) \quad \text{wobei } [\Theta^*] := [\{ W \otimes W \mid W \in L^* \}]$$

und dem Skalarprodukt $\langle A, B \rangle := \text{tr } AB$ für $A, B \in V^*$.

3.1.2. Existenz und Eindeutigkeit des GM-Schätzers

Lemma 3.1(3) Sei $\bar{W} \in L^* \cap S_+(V)$ so daß $\bar{W}V = \sum_{W \in L^*} WV$
 $\implies \bar{W} \otimes \bar{W} \in \bar{\Theta}^*$

Beweis: Es ist $\bar{Q}^* := \bar{W} \otimes \bar{W} \in [\Theta^*] \cap S_+(V^*)$.

Zu zeigen ist: $\sum_{Q^* \in \Theta^*} Q^* V^* \subseteq \bar{Q}^* V^*$,

bzw. äquivalent: $K(\bar{Q}^*) \subseteq \bigcap_{Q^* \in \Theta^*} K(Q^*)$

Sei $Y = y \otimes y \in K(\bar{Q}^*)$, d.h.

$$0 = \bar{Q}^* Y = \bar{W}_Y \otimes \bar{W}_Y$$

$$\implies y \in K(\bar{W}) = \bigcap_{W \in L^*} K(W)$$

$\implies Q^* Y = 0 \quad \forall Q^* \in \Theta^*$ nach Definition von Θ^* .

Wegen Corollar zu Lemma 3.1(1) folgt für beliebiges $A \in K(\bar{Q}^*)$: $Q^* A = 0 \quad \forall Q^* \in \Theta^*$.

Bemerkung: Falls etwa $L := \mathcal{L}(X)$, so kann man $\bar{W}$ wählen als $\bar{W} := X \Sigma X' + \sum c_i \cdot W_i$ mit $\Sigma = \Sigma' > 0$, $c_i > 0$.

Allgemein: $\bar{W} := P_L + \sum c_i \cdot W_i$ mit $c_i > 0$

(vergleiche Corollar 2 zu Theorem 1(5)),

denn wegen $P_L \in S(L) = L_1$ ist $L \subseteq \sum_{W \in L_1} WV$;

wegen $(1 \otimes 1)v = (1, v)1 \in L \quad \forall 1 \in L, v \in V$ ist mit Corollar zu Lemma 3.1(1) $\sum_{W \in L_1} WV \subseteq L$,

also insgesamt: $\bar{W}V = L + (\sum c_i W_i)V = \sum_{W \in L^*} WV$ und $\bar{W} \in L^* \cap S_+(V)$.

Lemma 3.1(4) Sei $\overline{Q^*} := \overline{W} \overline{\Theta} \overline{W} \in \overline{\Theta^*}$, $\overline{W}$ wie in Lemma 3.1(3)

$\implies \overline{Q^*}$ quasiregulär bezüglich L^*

Beweis: zu zeigen: $L^* \subseteq \overline{Q^*} V^*$

Sei $W \in L^*$ beliebig,

$Y := \overline{W}^+ W \overline{W}^+ = (\overline{Q^*})^+ W \in V^*$, so ist

$(\overline{Q^*}) Y = \overline{W} \overline{W}^+ W \overline{W}^+ \overline{W} = W \in (\overline{Q^*}) V^*$,

denn $\overline{W} \overline{W}^+ W = W \overline{W}^+ \overline{W} = W$ nach Wahl von $\overline{W}$.

Corollar: Sei $\overline{W}$ wie in Lemma 3.1(3) beliebig.

Es existiert ein GMO zum Modell $M(L^*, \Theta^*)$

$\langle \implies \rangle \quad A \overline{W}^+ B \overline{W}^+ A \in L^* \quad \forall A, B \in L^*$

Beweis: folgt aus Corollar 2 zu Lemma 2.1(5)
mit Lemma 3.1(4).

(vergleiche DRYGAS 1977, Theorem 4.6.,
RAO 1971 b), (6.6.))

Theorem 3(1) Sei $\overline{W}$ wie in Lemma 3.1(3) beliebig.

Es existiert ein GMO zum Modell $M(L^*, \Theta^*)$

$\langle \implies \rangle \quad A \overline{W}^+ A \in L^* \quad \neg A \in L^*$

Beweis: " $\Leftarrow$ " $A \overline{W}^+ A \in L^* \quad \forall A \in L^*$

$\implies A \overline{W}^+ B \overline{W}^+ A \in L^* \quad \neg A \in L^*$

(siehe DRYGAS 1977, Lemma 4.3.)

" $\Rightarrow$ " Sei $A \overline{W}^+ B \overline{W}^+ A \in L^* \quad \forall A \in L^*$,
dann ist insbesondere für

$B = \overline{W} \in L^*$:

$A \overline{W}^+ \overline{W} \overline{W}^+ A = A \overline{W}^+ A \in L^* \quad \neg A \in L^*$

(vergleiche DRYGAS 1977, Theorem 4.8. und
KLEFFE 1980, Theorem 2.2.)

3.1.3. Berechnung eines GM-Schätzers

Sei $L^* := L_1 + L_2$ mit $L_1 := S(L) = \{[1 \otimes 1 \mid 1 \in L]\}$,
 $L_2 := [W_1, \dots, W_m]$ mit $W_i = W_i' \geq 0$.

Sei $\bar{W} \in L^* \cap S_+(V)$ so daß $\bar{W}V = L + \sum W_i V$ gemäß Lemma 3.1(3),
 so ist $\bar{Q}^* := \bar{W} \otimes \bar{W} \in \bar{\Theta}^*$ und $\bar{Q}^*$ quasiregulär bezüglich L^* ,
 d.h. $K(\bar{W} \otimes \bar{W}) \subseteq (L^*)^\perp$.

Gesucht ist GY als Projektion von Y auf L^* entlang
 $(\bar{W} \otimes \bar{W}) (L^*)^\perp \oplus K(\bar{W} \otimes \bar{W})$,

d.h. $G: V^* \rightarrow L^*$ ist GMO zum Modell $M(L^*, \bar{Q}^*)$ in V^* .

Falls die Voraussetzungen von Theorem 3(1) erfüllt sind,
 gilt $M(L^*, \bar{Q}^*) \cong M(L^*, \Theta^*)$ und G ist GMO zum Modell $M(L^*, \Theta^*)$.

Sei $W_0: V \rightarrow L$ der Projektor auf L entlang $\bar{W}L^\perp \oplus K(\bar{W})$ in V,
 - z.B. $W_0 = P_L (P_L \bar{W}^+ P_L)^\perp P_L \bar{W}^+$ (siehe Theorem 1(5)) -

Lemma 3.1(5) $P_{L_1}^\circ := W_0 \otimes W_0$ ist Projektor auf L_1
 entlang $(\bar{W} \otimes \bar{W}) (L_1)^\perp \oplus K(\bar{W} \otimes \bar{W})$ in V^* .

Insbesondere ist $P_{L_1}^\circ$ GMO zum Modell $M(L_1, \bar{W} \otimes \bar{W})$.

Beweis:

- i) Für $l \in L$ ist $W_0 l = l$ und $P_{L_1}^\circ (1 \otimes 1) = 1 \otimes 1$,
 also $P_{L_1}^\circ \bar{W} = \bar{W} \quad \forall \bar{W} \in L_1$.
- ii) Für $y \in K(\bar{W}) \subseteq L^\perp$ ist $W_0 y = 0$ und $P_{L_1}^\circ (y \otimes y) = 0$,
 also $P_{L_1}^\circ \bar{W} = 0 \quad \forall \bar{W} \in K(\bar{W} \otimes \bar{W})$.
- iii) Für $Y = \bar{W} A \bar{W}$ mit $A \in (L_1)^\perp$, d.h. $\text{tr} A W = 0 \quad \forall W \in L_1$,
 ist $W_0 Y W_0' = W_0 \bar{W} A \bar{W} W_0' \in L_1$ und
 $0 = \text{tr} A W_0 \bar{W} A \bar{W} W_0' = 0$, also $W_0 \bar{W} A \bar{W} W_0' = 0$, da $W_0 \bar{W}$ symm.

Bemerkung: W_0 und damit $P_{L_1}^0$ hängen i.a. von dem speziell gewählten $\bar{W}$ ab.¹
 Falls jedoch $M(L, \Theta) \cong M(L, \bar{W})$, so ist W_0 GMO zu $M(L, \Theta)$ und unabhängig von dem speziell gewählten $\bar{W}$.

Corollar: $M(L, \bar{W}) \cong M(L, \Theta)$ in V

$$\iff M(L_1, \bar{W} \Theta \bar{W}) \cong M(L_1, \Theta^*) \text{ in } V^*$$

Beweis:

Sei W_0 GMO zu $M(L, \Theta) = M(L, \bar{W})$, so ist $W_0 W_i = W_i W_0'$
 für $i=1, \dots, m$.

Da für $A \in L_1 = S(L)$ $W_0 A = A$ gilt,
 folgt für beliebiges $W = A + \sum c_i W_i \in L^*$:
 $W_0 W = A + \sum c_i W_0 W_i$ symmetrisch.

$\implies (W_0 \Theta W_0') (W \Theta W) = W_0 W \Theta W W_0'$ symmetrisch $\forall W \in L^*$,
 also $(W_0 \Theta W_0') Q^*$ symmetrisch $\forall Q^* \in \Theta^*$.

Mit dem Symmetriekriterium (Lemma 0(4)) folgt die Behauptung.

Ist umgekehrt $(W_0 \Theta W_0') (W \Theta W)$ symmetrisch $\forall W \in L^*$,
 so ist $W_0 W_i$ symmetrisch $\forall i=1, \dots, m$
 und W_0 ist GMO zu $M(L, \Theta)$.

Setze $\tilde{W}_i := W_i - W_0 W_i W_0'$ ($i=1, \dots, m$), $\tilde{L}_2 := [\tilde{W}_1, \dots, \tilde{W}_m]$.

$\tilde{W}_i$ ($i=1, \dots, m$) und $\tilde{L}_2$ hängen i.a. von dem speziell gewählten $\bar{W}$ ab.

Falls jedoch $M(L, \bar{W}) \cong M(L, \Theta)$, so ist $\tilde{L}_2$ nur durch W_0 , also durch Θ bestimmt.

Wegen $W_0 W_i = W_i W_0'$ ($i=1, \dots, m$) ist dann $\tilde{W}_i = (I - W_0) W_i (I - W_0)'$.

Lemma 3.1(6) $L^* = L_1 \oplus \tilde{L}_2$ und

$$\forall W \in L_1, i=1, \dots, m \text{ ist } \text{tr } \bar{W} \bar{W}^+ \tilde{W}_i \bar{W}^+ = 0.$$

Beweis:

Offensichtlich ist $L^* = L_1 + L_2 = L_1 + \tilde{L}_2$.

Für $W \in L_1$ ist $W = W_0 W W_0^!$ und

$$\text{tr } \bar{W} \bar{W}^+ \tilde{W}_i \bar{W}^+ = \text{tr } \bar{W} \bar{W}^+ W_1 \bar{W}^+ - \text{tr } \bar{W} \bar{W}^+ W_0 W_1 W_0^! \bar{W}^+ = 0, \\ \text{denn } \bar{W}^+ W_0 = W_0^! \bar{W}^+.$$

Lemma 3.1(7) Es sei $P_{\tilde{L}_2}^0$ die Projektion auf $\tilde{L}_2$
entlang $(\bar{W} \otimes \bar{W}) (\tilde{L}_2)^\perp \oplus K(\bar{W} \otimes \bar{W})$,

$$\text{so gilt: } G = P_{L^*}^0 := P_{L_1}^0 + P_{\tilde{L}_2}^0.$$

Beweis:

i) $W \in L^*$ hat eine eindeutig bestimmte
Darstellung als $W = W_1 + W_2$ mit $W_1 \in L_1, W_2 \in \tilde{L}_2$.

$$P_{L^*}^0 W = P_{L_1}^0 W + P_{\tilde{L}_2}^0 W = W_1 + W_2 = W.$$

ii) Für $W \in K(\bar{W} \otimes \bar{W})$ ist $P_{L_1}^0 W = P_{\tilde{L}_2}^0 W = 0$,
also $P_{L^*}^0 W = 0$.

iii) Für $W \in (\bar{W} \otimes \bar{W}) (L^*)^\perp = (\bar{W} \otimes \bar{W}) ((L_1)^\perp \cap (L_2)^\perp)$
 $\subseteq (\bar{W} \otimes \bar{W}) (L_1)^\perp \cap (\bar{W} \otimes \bar{W}) (L_2)^\perp$
ist $P_{L_1}^0 W = P_{\tilde{L}_2}^0 W = 0$, also $P_{L^*}^0 W = 0$.

Lemma 3.1(8) Falls $\text{tr } \tilde{W}_i \bar{W}^+ \tilde{W}_j \bar{W}^+ = 0$ für $i \neq j$,
so hat $P_{\tilde{L}_2}^0$ die Darstellung:

$$P_{\tilde{L}_2}^0 Y = \sum \tilde{W}_i \cdot \frac{\text{tr } \bar{W}^+ \tilde{W}_i \bar{W}^+ Y}{\text{tr } \bar{W}^+ \tilde{W}_i \bar{W}^+ \tilde{W}_i} \\ = \sum \tilde{W}_i \cdot \frac{(\bar{W}^+ \tilde{W}_i \bar{W}^+ Y, Y)}{\text{tr } \bar{W}^+ \tilde{W}_i \bar{W}^+ \tilde{W}_i}.$$

(vergleiche KLEFFE 1976, (12), 1977, Theorem 7)

Beweis:

Durch $\langle A, B \rangle_0 := \text{tr } \bar{W}^+ A \bar{W}^+ B$ wird auf $(\bar{W} \otimes \bar{W}) V^*$ ein Skalarprodukt und auf V^* ein Semiskalarprodukt definiert.

Die Behauptung folgt aus Theorem 1(2)

(Aitken-Formel) mit Lemma 3.1(4) $(\bar{W} \otimes \bar{W})$ quasi-regulär bezüglich L^*).

Bemerkung: Falls die Bedingung " $\text{tr } \bar{W}_i \bar{W}^+ \bar{W}_j \bar{W}^+ = 0$ für $i \neq j$ " nicht erfüllt ist, muß man zuvor orthogonalisieren! (siehe Beispiele 3.3.2. und 3.3.3.)

Theorem 3(2) Sei $\bar{W} \in L^* \cap S_+(V)$ so daß $\bar{W}V = L + \sum W_i V$,

W_0 GMO zum Modell $M(L, \bar{W})$ in V und

$$\tilde{W}_i := W_i - W_0 W_i W_0' \quad (i=1, \dots, m).$$

Ferner sei $\text{tr } \bar{W}_i \bar{W}^+ \bar{W}_j \bar{W}^+ = 0$ für $i \neq j$.

So gilt:

$$i) \quad L^* = L_1 \oplus \tilde{L}_2 \quad \text{mit } \tilde{L}_2 := [\tilde{W}_1, \dots, \tilde{W}_m]^\perp.$$

$E_P Y$ hat eine eindeutig bestimmte Darstellung als $E_P Y = W + \sum c_i \tilde{W}_i$ mit $W \in L_1, c_1, \dots, c_m \in \mathbb{R}$.

(W und $c_1, \dots, c_m$ hängen i.a. von $\bar{W}$ ab.)

$$ii) \quad \text{Für } \hat{c}_i := \frac{(\bar{W}^+ \tilde{W}_i \bar{W}^+ Y, Y)}{\text{tr } \bar{W}^+ \tilde{W}_i \bar{W}^+ \tilde{W}_i} \quad \text{gilt:}$$

$\hat{c}_i$ ist erwartungstreuer quadratischer Schätzer für c_i ($i=1, \dots, m$).

In der Klasse aller erwartungstreuen quadratischen Schätzer für c_i ist $\hat{c}_i$ lokal optimal für $E_P Y = 0$ und $\text{Cov}_P Y = \sum c_i \tilde{W}_i = \bar{W}$.

- iii) Falls $\overline{A} \overline{W}^+ A \in L^*$ $\forall A \in L^*$, so ist $\hat{c}_i$ ($i=1, \dots, m$) unabhängig von dem speziell gewählten $\overline{W}$ und gleichmäßig optimal $\forall E_P Y \in L$ und $\text{Cov}_P Y \in \Theta$.

Beweis:

- ii) Für $E_P Y = \mu$, $W := \mu \mu'$, und $\text{Cov}_P Y = \sum c_i W_i$ ist $E_P Y = W + \sum c_i W_i$ und $\text{Cov}_P Y = 2((W + \sum c_i W_i) \otimes (W + \sum c_i W_i) - W \otimes W)$.
 $\hat{c}_i$ ($i=1, \dots, m$) ist optimal, falls $\text{Cov}_P Y = \overline{W} \otimes \overline{W}$, also $W = 0$ und $\sum c_i W_i = \overline{W}$.

- iii) Falls $\overline{A} \overline{W}^+ A \in L^*$ $\forall A \in L^*$, so ist gemäß Theorem 3(1) $P_{L^*}^\circ = P_{L_1}^\circ + P_{L_2}^\circ$ GMO zu $M(L^*, Q^*)$ $\forall Q^* \in \Theta^*$, d.h. $P_{L^*}^\circ Y$ ist gleichmäßig optimal $\forall E_P Y \in L$ und $\text{Cov}_P Y \in \Theta$.
Außerdem hängt $P_{L^*}^\circ$ nicht von dem speziell gewählten $\overline{W}$ ab.

Ich zeige: $\hat{c}_i$ ($i=1, \dots, m$) hängt nicht von $\overline{W}$ ab:

$\overline{W}$ und $\overline{\overline{W}}$ mögen die Voraussetzungen von Lemma 3.1(3) erfüllen;

$\overline{W}$ bestimme eine Zerlegung von L^* in

$L_1 \oplus [W_1, \dots, W_m]$ gemäß i),

$\overline{\overline{W}}$ bestimme eine Zerlegung in $L_1 \oplus [\overline{W}_1, \dots, \overline{W}_m]$.

Ferner sei $P_{L^*}^\circ Y = W + \sum c_i W_i = \overline{W} + \sum \tilde{c}_i \overline{W}_i$ mit $W, \overline{W} \in L_1$ und $c_i, \tilde{c}_i$ ($i=1, \dots, m$) $\in \mathbb{R}$.

Sei W_0 GMO zu $M(L, \overline{\overline{W}})$ in V , so gilt andererseits $W + \sum c_i W_i = W + W_0 (\sum c_i W_i) W_0' + \sum c_i (W_i - W_0 W_i W_0')$

so daß wegen $\overline{W}_i = W_i - W_0 W_i W_0'$ ($i=1, \dots, m$) $\overline{W} = W + W_0 (\sum c_i W_i) W_0' \in L_1$ und $\tilde{c}_i = c_i$ $\forall i \leq m$.

Corollar: Falls $\overline{A} \overline{W}^+ A \in L^* \quad \forall A \in L^*$, so gilt für $i=1, \dots, m$:

- i) $\hat{c}_i \geq 0$
- ii) $\hat{c}_i$ ist invariant gegen Translationen von $E_P Y$.

Beweis:

Für $i=1, \dots, m$ hängt $\hat{c}_i$ nicht von dem speziell gewählten $\overline{W}$ ab,

insbesondere kann man für jedes $i \in \{1, \dots, m\}$ $\overline{W}$ "geeignet" wählen:

Sei $i \in \{1, \dots, m\}$ beliebig, fest gewählt,

P_i die Orthogonalprojektion auf $K(W_i) \cap \sum_{W \in L^*} W V$,

so daß $\overline{W} := W_i + P_i$ die Voraussetzungen von Lemma 3.1(3) erfüllt.

Sei ferner W_0 GMO zu $M(L, \overline{W}) \cong M(L, W_i)$ und

$$\tilde{W}_i := W_i - W_0 W_i W_0'.$$

Da nach Wahl von W_0 $W_0 W_i = W_i W_0'$, folgt

$$\tilde{W}_i = (I - W_0) W_i (I - W_0)' \geq 0 \quad \text{und} \quad \overline{W}^+ \tilde{W}_i \overline{W}^+ \geq 0,$$

also $\hat{c}_i \geq 0$.

Wegen $(\overline{W}^+ \tilde{W}_i \overline{W}^+ 1, 1) = 0 \quad \forall 1 \in L$ gemäß Lemma 3.1(6)

folgt: $\overline{W}^+ \tilde{W}_i \overline{W}^+ 1 = 0 \quad \forall 1 \in L$.

(vergleiche KLEFFE & PINCUS 1974, Theorem 6)

3.1.4. Invarianz und Positivität des GM-Schätzers

Lemma 3.1(9) Sei $\bar{W} \in L^* \cap S_+(V)$ so daß $\bar{W}V = L + \sum W_i V$
 ($\bar{W}$ fest gewählt gemäß Lemma 3.1(3))

W_0 GMO zum Modell $M(L, \bar{W})$ wie in Kapitel 1.
 (W_0 Projektor auf L entlang $\bar{W}L^\perp \oplus (L^\perp \cap K(\bar{W}))$),
 $\hat{c}_i$ ($i=1, \dots, m$) wie in Theorem 3(2).

Falls $M(L, 0) \cong M(L, \bar{W})$, so gilt:

- i) $\hat{c}_i \geq 0$ ($i=1, \dots, m$)
- ii) $\hat{c}_i$ ist invariant gegen Translationen von $E_P Y$ in L ($i=1, \dots, m$).

iii) $\hat{c}_i$ hat die Darstellung

$$\hat{c}_i = \frac{(\bar{W}^+ \bar{W}_i \bar{W}^+ (I - W_0) Y, (I - W_0) Y)}{\text{tr } \bar{W}^+ \bar{W}_i \bar{W}^+ \bar{W}_i} \quad (i=1, \dots, m),$$

wobei $\bar{W} := (I - W_0) \bar{W} (I - W_0)' = \sum a_i \bar{W}_i$ mit $a_i \in \mathbb{R}$.

- iv) $\hat{c}_i$ ist lokal optimal, falls $c_i = a_i$ ($i=1, \dots, m$)
 - d.h. falls $\text{Cov}_{PY} = \sum a_i W_i$ - und $E_P Y \in L$ bel.

Beweis:

- i) W_0 GMO zu $M(L, 0) \Rightarrow W_0 W_i = W_i W_0' \quad \forall i=1, \dots, m$
 und $\bar{W}_i = (I - W_0) W_i (I - W_0)'$.

Wegen $\bar{W}_i \geq 0$, folgt $\hat{c}_i \geq 0 \quad \forall i=1, \dots, m$.

- ii) Wegen Lemma 3.1(6) ist für beliebiges $W \in L_1$
 $\text{tr } \bar{W}^+ \bar{W}_i \bar{W}^+ W = 0 \quad \forall i=1, \dots, m$;
 insbesondere ist $\forall l \in L \quad (\bar{W}^+ \bar{W}_i \bar{W}^+ l, l) = 0$,
 so daß $\bar{W}^+ \bar{W}_i \bar{W}^+ l = 0 \quad \forall l \in L$.

Corollar: Falls $M(L, \theta) \cong M(L, \bar{W})$, so sind die $\hat{c}_i$ ($i=1, \dots, m$) gerade die von Kleffe angegebenen MINQUE-Schätzer für c_i .
(siehe KLEFFE 1976, Lemma 3, bzw. 1980, Theorem 1)

Bemerkung: Falls zu $M(L, \theta)$ ein GMO existiert, so hängen zwar die $c_1, \dots, c_m$ sowie die $\bar{W}_1, \dots, \bar{W}_m$ nicht von $\bar{W}$ bzw. $\bar{W}$ ab, die Schätzer $\hat{c}_1, \dots, \hat{c}_m$ i.a. jedoch sehr wohl.
Nur falls $\bar{A}\bar{W}^+A \in L^* \quad \forall A \in L^*$ (siehe Theorem 3(2)), sind die $\hat{c}_i$ unabhängig von dem speziell gewählten $\bar{W}$ und gleichmäßig optimal $\forall c_1, \dots, c_m \in \mathbb{R}$.

Lemma 3.1(10) Sei W_0 ein beliebiger Projektor auf L ,

$$\tilde{W}_i := (I - W_0)W_i(I - W_0)' \quad \text{für } i=1, \dots, m$$

$$\bar{\bar{W}} := \sum a_i \tilde{W}_i \quad \text{mit } a_1, \dots, a_m \geq 0 \text{ so daß } \sum a_i W_i V = \sum W_i V.$$

Ferner sei $L^\perp \subseteq \sum W_i V$

(das ist z.B. erfüllt, falls $\bigcap K(W_i) = 0$).

$$\text{Für } \hat{c}_i := \frac{(\bar{\bar{W}}^+ \tilde{W}_i \bar{\bar{W}}^+ (I - W_0) y, (I - W_0) y)}{\text{tr } \bar{\bar{W}}^+ \tilde{W}_i \bar{\bar{W}}^+ \tilde{W}_i} \quad \text{gilt:}$$

- i) $\hat{c}_i$ ($i=1, \dots, m$) hängt nicht von dem speziell gewählten Projektor W_0 ab, insbesondere kann man $W_0 = P_L$ setzen.
- ii) $\hat{c}_i$ ($i=1, \dots, m$) sind gerade die (bei $\text{Cov}_P y = \sum a_i W_i$ lokal) besten invarianten Schätzer für c_i .

Beweis:

i) ich zeige:

$$(I - W_0)' \bar{\bar{W}}^+ (I - W_0) = [(I - P_L) \sum a_i W_i (I - P_L)]^+$$

Sei $W := \sum a_i W_i$.

$V = L \oplus WL^\perp \oplus (L^\perp \cap K(W)) = L \oplus WL^\perp$ nach Vor.

Für $l \in L$ ist $(I-W_0)l=l$ und $(I-W_0)'l=0$, so daß
 $(I-W_0)' \bar{W}^+ (I-W_0)l = 0$.

Ebenso ist $[(I-P_L)W(I-P_L)]^+ l = 0$.

Für $y = Wa$ mit $a \in L^\perp$ ist

$$(I-W_0)' [(I-W_0)W(I-W_0)']^+ (I-W_0)Wa = a,$$

da $(I-W_0)' [(I-W_0)W(I-W_0)']^+ (I-W_0)W$

Projektor auf $(I-W_0)' V = L^\perp$ ist.

(Aitken-Formel, W^+ quasiregulär bzgl. $L^\perp$)

$$\begin{aligned} [(I-P_L)W(I-P_L)]^+ Wa &= \\ &= (I-P_L) [(I-P_L)W(I-P_L)]^+ (I-P_L)Wa = a. \end{aligned}$$

ii) siehe KLEFFE 1976, Theorem 1

Theorem 3(3) Sei $L^\perp \subseteq \sum W_i V$.

$\bar{W} := W + \sum a_i W_i$ mit $W \in L_1 \cap S_+(V)$, $a_1, \dots, a_m \geq 0$
 so daß $\bar{W}V = L + \sum W_i V$,

W_0 GMO zu $M(L, \bar{W})$ in V ,

$\hat{W}_i := W_i - W_0 W_i W_0'$ und $\tilde{W}_i := (I-W_0)W_i(I-W_0)'$
 für $i=1, \dots, m$

$$\hat{c}_i := \frac{(\bar{W}^+ \hat{W}_i \bar{W}^+ y, y)}{\text{tr } \bar{W}^+ \hat{W}_i \bar{W}^+ \hat{W}_i} \quad (i=1, \dots, m) \quad \text{sowie}$$

$$\tilde{c}_i := \frac{(\bar{W}^+ \tilde{W}_i \bar{W}^+ y, y)}{\text{tr } \bar{W}^+ \tilde{W}_i \bar{W}^+ \tilde{W}_i} \quad (i=1, \dots, m),$$

so gilt:

- i) $\hat{c}_i$ ist (bei $E_p y = 0$, $\text{Cov}_p y = \sum a_i W_i$ lokal) bester erwartungstreuer quadratischer Schätzer für c_i ($i=1, \dots, m$).
- ii) $\hat{\bar{c}}_i$ ist (bei $\text{Cov}_p y = \sum a_i W_i$ lokal) bester invarianter erwartungstreuer quadratischer Schätzer für c_i ($i=1, \dots, m$).
- iii) $\hat{c}_i = \hat{\bar{c}}_i \quad \forall i=1, \dots, m, \quad \forall y \in V$
 $\Longleftrightarrow M(L, \Theta) \cong M(L, \bar{W})$

Beweis:

- i) siehe Theorem 3(2)
- ii) nach Wahl von W_0 ist $W_0 \bar{W} = \bar{W} W_0'$ und
 $\bar{W} = W_0 \bar{W} W_0' + (I - W_0) \bar{W} (I - W_0)'$
 $= W_0 \bar{W} W_0' + \sum a_i \tilde{W}_i = W_0 \bar{W} W_0' + \bar{W}$
 $(\bar{W} \text{ wie in Lemma 3.1(10)})$
 $\bar{W}^+ = W_0' (W_0 \bar{W} W_0')^+ W_0 + (I - W_0)' \bar{W}^+ (I - W_0)$
 $(\text{siehe Beweis zu Lemma 3.1(9)})$
 $\bar{W}^+ \tilde{W}_1 \bar{W}^+ = (I - W_0)' \bar{W}^+ \tilde{W}_1 \bar{W}^+ (I - W_0).$
 Mit Lemma 3.1(10) folgt die Behauptung.

- iii) $M(L, \Theta) \cong M(L, \bar{W}) \implies \tilde{W}_i = \tilde{\bar{W}}_i \quad (i=1, \dots, m)$
 und $\hat{c}_i = \hat{\bar{c}}_i \quad (i=1, \dots, m)$ (siehe Theorem 3(2)ii)).

Sei umgekehrt $\hat{c}_i = \hat{\bar{c}}_i \quad \forall i=1, \dots, m, \quad \forall y \in V$.

Insbesondere für $y \in L + \sum W_i V$:

(für $y \in (L + \sum W_i V)^\perp = L^\perp \cap K(W_1)$ ist immer $\hat{c}_i = \hat{\bar{c}}_i = 0$, $i=1, \dots, m$)

$$\hat{c}_i = \langle \tilde{W}_1, y \otimes y \rangle_0 = \langle \tilde{\bar{W}}_1, y \otimes y \rangle_0 = \hat{\bar{c}}_i \quad (i=1, \dots, m)$$

($\langle A, B \rangle_0 := \text{tr } A \bar{W}^+ B \bar{W}^+$ definiert ein Semiskalarprodukt auf V^*)

$$\Rightarrow \tilde{W}_i = \tilde{W}_i \quad (i=1, \dots, m),$$

$$W_i = (I - W_0)W_i(I - W_0)' + W_0W_iW_0'. \quad (*)$$

Ich zeige: $L \cap \sum W_i L^\perp = 0$.

Sei $l = \sum W_i b_i$ mit $b_i \in L^\perp$, $l \in L$,

Wegen (*) ist mit $(I - W_0)' b_i = b_i$, $W_0' b_i = 0$

$$l = \sum (I - W_0)W_i b_i = (I - W_0) \sum W_i b_i = (I - W_0)l = 0;$$

d.h. zu $M(L, \Theta)$ existiert ein GMO,

$$M(L, \Theta) \cong M(L, \bar{W}) \cong M(L, \sum a_i W_i).$$

Corollar: $M(L^*, \Theta^*) \cong M(L^*, \bar{W} \otimes \bar{W}) \Rightarrow M(L, \Theta) \cong M(L, \bar{W})$

Beweis:

Falls $M(L^*, \Theta^*) \cong M(L^*, \bar{W} \otimes \bar{W})$, so folgt mit

Corollar zu Lemma 3(2) die Invarianz der $\hat{C}_i$ ($i=1, \dots, m$),

also $\hat{C}_i = \hat{C}_i$ ($i=1, \dots, m$).

Mit Theorem 3(3) iii) folgt die Behauptung.

3.2. Missing-Data-Problem - Das Modell $M(P^*L^*, P^*\odot P^*)$

Es seien die Voraussetzungen von Kapitel 1.2.0. erfüllt.

Statt $y = (y_1, \dots, y_n)' \in V$ werde nur $P_y = (\Delta_1 y_1, \dots, \Delta_n y_n)' \in PV$ beobachtet, wobei $P = P' = P^2 = \text{diag} (\Delta_1, \dots, \Delta_n) \in S_+(V)$ ein nicht-stochastischer Operator sei.

Wegen $P_y \odot P_y = (P \odot P)Y$ wird durch $P^* := P \odot P$ auf V^* ein Projektor definiert.

Entsprechend Kapitel 1.2.0. ist als neues Modell $M(P^*L^*, P^*\odot P^*)$ in P^*V^* bzw. in V^* zu betrachten, wobei

$$P^*L^* = \{PWP \mid W \in L^*\} \quad \text{und}$$

$$P^*\odot P^* = \left[\{ PWP \odot PWP \mid W \in L^* \} \right] \cap S_+(V^*) \quad \text{ist.}$$

Auf das Modell $M(P^*L^*, P^*\odot P^*)$ sind die Sätze über Existenz und Berechnungsmöglichkeiten lokal und gleichmäßig bester Schätzer gemäß Abschnitt 3.1. anzuwenden.

Die Anwendung von Theorem 1(7) bzw. 1(8) über das "Ersetzen fehlender Daten" auf $M(P^*L^*, P^*\odot P^*)$ bzw. $M(P^*L^*, P^*(\bar{W} \odot \bar{W})P^*)$ - $\bar{W}$ fest gewählt - liefert i.a. keine Möglichkeit, den Beobachtungsvektor P_y zu ergänzen, sondern nur (falls die Voraussetzungen erfüllt sind) das Ergänzen von P^*Y durch eine Matrix $B := AP^*Y$, für die $B = (I - P \odot P)P_{L^*}^{\circ}(PYP + B)$ gelten muß.

Die Beispiele 3.3.1. - 3.3.3. verdeutlichen, wie man auf diese Weise die neuen Schätzfunktionen berechnen kann.

In einem wichtigen Spezialfall ist es jedoch möglich, die bekannten Schätzfunktionen (lediglich mit einem Korrekturfaktor versehen) auf den "erweiterten Beobachtungsvektor" $Py + x_0$ anzuwenden, wobei x_0 als Schätzwert für die fehlenden Meßwerte $(I-P)y$ angesehen werden kann.

Lemma 3.2(1) Sei $I \in \Theta$, $\text{tr} (W_i (P - P_{PL}) W_j (P - P_{PL})) = 0$ für $i \neq j$,
 $P = P' = P^2$ so daß $\dim PL = \dim L$.

So gilt:

Die (bei $\text{Cov}_P y = \sigma^2 I - \sigma^2 > 0$ beliebig - lokal) besten invarianten erwartungstreuen quadratischen Schätzer für c_i ($i=1, \dots, m$) sind zu berechnen aus:

$$\hat{c}_i = \frac{(W_i (I - P_L) (Py + x_0), (I - P_L) (Py + x_0))}{\text{tr} [(P - P_{PL}) W_i]^2},$$

wobei x_0 die eindeutig bestimmte Lösung von
 $x = (I - P) P_L (Py + x)$ ist.

Beweis:

Wegen Lemma 3.1(10) mit $\bar{W} := I$, $\tilde{W}_i := (P - P_{PL}) W_i (P - P_{PL})$ ist

$$\hat{c}_i = \frac{(P W_i P (I - P_{PL}) P y, (I - P_{PL}) P y)}{\text{tr} [(P - P_{PL}) W_i]^2}.$$

Mit Theorem 1(7) folgt für $P(I - P_{PL})Py$

$$\begin{aligned} P(I - P_{PL})Py &= Py - P P_L (Py + x_0) \\ &= Py - P_L (Py + x_0) + (I - P) P_L (Py + x_0) \\ &= Py + x_0 - P_L (Py + x_0) = (I - P_L) (Py + x_0). \end{aligned}$$

Corollar: Es seien die Voraussetzungen von Lemma 3.2(1) erfüllt.

Die $\tilde{c}_i$ ($i=1, \dots, m$) aus Lemma 3.2(1) sind genau dann die besten erwartungstreuen quadratischen Schätzer für c_i , wenn $M(PL, P \odot P) \cong M(PL, P)$.

Beweis: folgt aus Theorem 3(3) mit $\bar{W} := I$ und Lemma 3.2(1).

Bemerkung: Unter den Voraussetzungen von Lemma 3.2(1) sind die Schätzfunktionen $\tilde{c}_i$ ($i=1, \dots, m$) im Wesentlichen die Schätzfunktionen $\hat{c}_i$ ($i=1, \dots, m$) für das vollständige Modell $M(L, I)$, angewendet auf $Py + x_0$. Zu beachten ist, daß i.a. $\text{tr}[(P - P_{PL})W_i]^2 < \text{tr}[(I - P_L)W_i]^2$ ist.

$$\tilde{c}_i(Py) = \frac{\text{tr}[(I - P_L)W_i]^2}{\text{tr}[(P - P_{PL})W_i]^2} \hat{c}_i(Py + x_0).$$

(Siehe Beispiele 3.3.1. und 4.3.)

Anmerkungen zu 3.1. und 3.2.

- 1) Drygas hat zum quadratischen Schätzen von Varianzkomponenten bereits ein lineares Modell in $Y:=yy'$ betrachtet (DRYGAS 1976, Beispiel 2.2., DRYGAS 1977).

Ein ähnlicher Ansatz findet sich bei Puckelsheim (PUCKELSHEIM 1976, 1977).

Das von mir eingeführte Konzept der Äquivalenzrelation zwischen linearen Modellen (siehe Lemma 0(5)) in Verbindung mit der sehr einfachen Darstellung von $[\theta^*]$ (siehe Abschnitt 3.1.1.) ermöglichen es, noch konsequenter Sätze aus der Theorie linearer Schätzer zu übertragen. DRYGAS 1977, Theorem 4.6. und 4.8. sowie SEELY 1971, Theorem 1 und Theorem 3 ergeben sich unmittelbar als Spezialfälle (siehe Corollar zu Lemma 3.1(4) und Theorem 3(1)).

Im Übrigen habe ich in meinen Ausführungen von der speziellen Gestalt des "Beobachtungsvektors" als reeller symmetrischer Matrix abgesehen und Vektoren im abstrakten Tensorraum $V \otimes V$ betrachtet, d.h. auch zur Behandlung quadratischer Schätzer "koordinatenfrei" argumentiert.

- 2) Im Zusammenhang mit der Berechnung (lokal oder gleichmäßig) bester Schätzer wird der Vorteil der geometrischen Betrachtungsweise besonders deutlich:

Die algebraische Herleitung des entsprechenden Projektors ist offensichtlich direkter und anschaulicher als das Minimieren der zugehörigen Varianzen (siehe z.B. RAO 1971 b), Abschnitt 3).

Außerdem wird die enge Verbindung zur "MINQUE-Theorie" (unter Normalverteilungsannahme) offensichtlich:

Die Beliebigkeit in der Wahl der Norm entspricht hier dem durch die tatsächliche Kovarianz der Beobachtungen bestimmten Skalarprodukt.

Dieser Zusammenhang wird auch von Kleffe beschrieben (KLEFFE 1976, insbesondere Theorem 1).

- 3) In der gesamten mir bekannten Literatur über das quadratische Schätzen von Varianz- und Kovarianzkomponenten (mit der einen unten genannten Ausnahme) werden die Fragen nach der Existenz (gleichmäßig) bester quadratischer Schätzer und (gleichmäßig) bester invarianter quadratischer Schätzer als zwei voneinander unabhängige Probleme betrachtet.

Insbesondere Drygas (DRYGAS 1977) gibt notwendige und hinreichende Bedingungen für die Existenz beider Typen von Schätzern an, die eine deutliche strukturelle Ähnlichkeit aufweisen (DRYGAS 1977, Theorem (4.8) sowie Theorem (5.6) und (5.8)). Ungeklärt bleibt jedoch die Frage nach dem mathematischen Zusammenhang zwischen beiden Schätzern. Die Entscheidungskriterien für einen der beiden Typen sind ausschließlich im Vorfeld der mathematischen Analyse zu suchen.

Nun ist jedoch für bestimmte balanzierte Versuchsanordnungen bekannt, daß die (existierenden) gleichmäßig besten quadratischen erwartungstreuen Schätzer invariant gegen Translationen von $E_{\mathbf{y}}$ sind, daß also beide Schätzer identisch sind.

Kleffe und Rincus haben für den Spezialfall " $\mathbf{I} \in \Theta$ " auf dem Umweg über Bayes-Schätzer nachgewiesen, daß allgemein im Falle der Existenz gleichmäßig bester quadratischer erwartungstreuer Schätzer diese bereits invariant sind.

(KLEFFE & PINCUS 1974, Theorem 6).

Ich konnte darüberhinaus zeigen (Theorem 3(3)), daß auch lokal beste quadratische Schätzer u.U. diese Invarianzeigenschaft besitzen, nämlich genau dann wenn $E_{\mathbf{y}} \quad \forall \mathbf{W} \in \Theta$ gleichmäßig optimal linear schätzbar ist.

Dieses Ergebnis hat eine gewisse anschauliche Plausibilität:
Wenn E_{PY} $\forall W \in \mathcal{O}$ gleichmäßig optimal schätzbar ist, so hängen die Schätzwerte für c_i ($W = \sum c_i W_i = \text{Cov}_P Y$) nicht von dem tatsächlichen Wert E_{PY} ab.

Dieser Zusammenhang ist jedoch als Folge der Normalverteilungsannahme bzw. der Konstruktion von $M(L^*, \theta^*)$ aus $M(L, \theta)$ zu verstehen. Unter allgemeineren Annahmen über $\text{Cov}_P Y$ ist ein derartiges Ergebnis vermutlich nicht zu bestätigen.

3.3. Beispiele

3.3.0. Regressionsmodell mit festen und zufälligen Effekten

Sei V, L, Θ wie in 2.3.0.,

$$V^* := S(V), L^* := L_1 + L_2 + [I_{nm}] \quad \text{mit } L_1 := \{X_1 \Sigma X_1' \mid \Sigma = \Sigma' \geq 0\}$$

$$L_2 := \{X_2 \Sigma X_2' \mid \Sigma = \Sigma' \geq 0, \dots\},$$

X_1 und X_2 bekannte reelle Matrizen

$$[\Theta^*] := [\{W \otimes W \mid W \in L^*\}]$$

Wegen $I \in L^*$ existiert ein GMO zum Modell $M(L^*, \Theta^*)$ in V^* genau dann wenn $W^2 \in L^* \quad \forall W \in L^*$ (Theorem 3(1) mit $\bar{W} = I$).

Behauptung: Falls $\mathcal{N}(X_2) \subseteq \mathcal{N}(X_1)$, so ist P_{L^*} GMO zum Modell $M(L^*, \Theta^*)$.

Beweis: $\mathcal{N}(X_2) \subseteq \mathcal{N}(X_1) \implies$

$$L_2 \subseteq S(\mathcal{N}(X_2)) \subseteq S(\mathcal{N}(X_1)) = L_1 \quad \text{und}$$

$$L^* = L_1 \oplus [I].$$

Für $W = X_1 \Sigma X_1' + cI \in L^*$ ist dann

$$W^2 = X_1 (\Sigma X_1' X_1 \Sigma + 2c\Sigma) X_1' + c^2 I \in L^*.$$

Behauptung: Sei $P = P' = P^2 \in S_+(V)$ beliebig, $P^* := P \otimes P$ und $\mathcal{N}(X_2) \subseteq \mathcal{N}(X_1)$, so gilt:

$P_{P^* L^*}$ ist GMO zum Modell $M(P^* L^*, P^* \Theta^* P^*)$.

Beweis: Wegen Theorem 3(1) mit $\bar{W} = I$ ist zu zeigen:

$$P W P W P \in P^* L^* \quad \forall W \in L^*.$$

Sei $W = X_1 \Sigma X_1' + cI \in L^*$, so ist

$$P W P W P = P X_1 (\Sigma X_1' P X_1 \Sigma + 2c \Sigma) X_1' P + c^2 P$$

$$\in P^* L^*.$$

Bemerkung 1: Falls $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$, so sind offenbar einzelne "Varianzkomponenten" - außer der Fehlervarianz σ_e^2 - nicht schätzbar:

$$\begin{aligned} \text{Sei etwa } E_P Y &= X_1 \Sigma_1 X_1' + X_2 \Sigma_2 X_2' + cI \\ &= X_1 \Sigma X_1' + cI \end{aligned}$$

$$(*) \quad \text{mit } \Sigma = \Sigma_1 + (X_1' X_1)^{-1} X_1' X_2 \Sigma_2 X_2' X_1 (X_1' X_1)^{-1}$$

so ist $P_{L^*} Y =: X_1 \hat{\Sigma} X_1' + \hat{c}I$, wobei $\hat{\Sigma}$ und $\hat{c}$ als Schätzungen für Σ bzw. c anzusehen sind.

Σ_1 und Σ_2 sind durch (*) nicht eindeutig bestimmt.

Bemerkung 2: Falls $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$, so ist $M(L, \emptyset) \cong M(L, I)$ und $M(PL, P \emptyset P) \cong M(PL, P) \quad \forall P = P' = P^2$ (siehe Abschnitt 2.3.0.).

Wegen Lemma 3.1(9) sind die besten quadratischen Schätzer für die Fehlervarianz σ_e^2 jeweils invariant gegen Translationen von $E_P Y$ in L bzw. von $E_P P Y$ in PL .

Im Falle unvollständiger Daten ist wegen Lemma 3.2(1) zur Berechnung von $\tilde{\sigma}_e^2$ der Beobachtungsvektor $P Y$ "durch Schätzwerte zu ergänzen". (siehe Beispiel 3.3.1.)

3.3.1. $\mathcal{M}(X_2) \subseteq \mathcal{M}(X_1)$: BQUEs für σ_e^2

Beispiel: Einfache Klassifikation

Sei $V := \mathbb{R}^{nm}$, $L := \mathcal{M}(X_1)$

$$\Theta := \{ \sigma_a^2 X_2 X_2' + \sigma_e^2 \cdot I_{nm} \mid \sigma_a^2, \sigma_e^2 \geq 0 \} = \{ \sigma_a^2 (I_n \otimes E_m) + \sigma_e^2 \cdot I_{nm} \}$$

$$\text{mit } X_1 := \begin{pmatrix} 1 & A_1 \\ \vdots & \vdots \\ 1 & A_n \end{pmatrix} \quad X_2 := \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \quad A_i \text{ wie in 1.3.1.}$$

Äquivalente Formulierung:

$$\begin{aligned} Y_{ij} &= \mu + \alpha_i + a_i + e_{ij} \quad i=1, \dots, n; j=1, \dots, m; \\ E_P Y_{ij} &= \mu + \alpha_i \quad \text{und} \quad \text{Cov}_P(a_1, \dots, a_n)' = \sigma_a^2 \cdot I_n \\ &\quad \text{Cov}_P(e_{11}, \dots, e_{nm})' = \sigma_e^2 \cdot I_{nm} \end{aligned}$$

Quadratisches Modell:

$$V^* := S(\mathbb{R}^{nm}),$$

$$L^* := \{ X_1 \Sigma X_1' \mid \Sigma = \Sigma' \geq 0 \} + [I_{nm}] = L_1 + [I_{nm}]$$

$$[\Theta^*] := [W \otimes W \mid W \in L^*]$$

Berechnung von P_{L^*} :

Setze $W_0 := X_1 (X_1' X_1)^{-1} X_1' = \frac{1}{m} (I_n \otimes E_m)$, so ist $I - W_0 \in L^*$

und $\text{tr } W(I - W_0) = 0 \quad \forall W \in L_1$,

$$\text{d.h. } L^* = L_1 \oplus [I - W_0] \quad \text{und} \quad P_{L^*} = P_{L_1} + P_{[I - W_0]}.$$

$$\begin{aligned}
 P_{L^*} Y &= (W_0 \otimes W_0) Y + (I - W_0) \frac{\langle (I - W_0), Y \rangle}{\langle (I - W_0)(I - W_0) \rangle} \\
 &= (W_0 Y \otimes W_0 Y) + (I - W_0) \frac{((I - W_0) Y, Y)}{\text{tr}(I - W_0)} \\
 &= \left[(\bar{Y}_i, \bar{Y}_{1.})_{i, l=1, \dots, n} \otimes E_m \right] + \\
 &\quad + (I - W_0) \frac{\sum y_{ij} - \frac{1}{m} \sum_i y_i^2}{n(m-1)}.
 \end{aligned}$$

Als gleichmäßig bester Schätzer für σ_e^2 ergibt sich somit

$$\hat{\sigma}_e^2 = \frac{1}{n(m-1)} \left(\sum y_{ij}^2 - \frac{1}{m} \sum_i y_i^2 \right)$$

(vergleiche DRYGAS 1977, (7.1) und 1972, (1.4))

Missing-Data-Problem

Sei $P := \text{diag} (I_{nm-1}, 0)$ wie in 1.3. und 2.3.,
d.h. Beobachtung y_{nm} fehle, und $P^* := P \otimes P$.

Wegen $\mathcal{M}(X_2) = \mathcal{M}(X_1)$ ist $P_{P^*L^*}$ GMO zum Modell $M(P^*L^*, P^*O^*P^*)$
und wegen $\dim P^*L^* = \dim L^* = nm + 1$ ist

$P_{P^*L^*} Y = P^* P_{L^*} (PYP + B)$, wobei $B \in (I - P \otimes P)L^*$ die Gestalt

$$\left(\begin{array}{c|c} 0_{nm-1} & \begin{array}{c} c_1 \\ \vdots \\ c_1 \\ c_2 \\ \vdots \\ c_n \\ \vdots \\ c_n \end{array} \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} m \text{ Zeilen} \\ \\ m-1 \text{ Zeilen} \end{array}$$

$$c_1 \dots c_1 \quad c_2 \dots c_n \quad c_n \quad c_n + c_0$$

mit $c_1, \dots, c_n, c_n + c_0 \in \mathbb{R}$ hat. $c_1, \dots, c_n, c_0$ sind als eindeutig
bestimmte Lösungen des linearen Gleichungssystems

(*) $B = (I - P \otimes P) P_{L^*} (PYP + B)$ zu bestimmen. (Theorem 1(7))

Es ist $P_{L^*} (P^*Y + B) = W_0 P_Y \otimes W_0 P_Y + W_0 B W_0 +$

$$+ \frac{(I - W_0)}{n(m-1)} \left\{ ((I - W_0) P_Y, P_Y) + \text{tr} (I - W_0) B \right\}$$

Als gleichmäßig bester Schätzer für σ_e^2 im Modell $M(P^*L^*, P^*O^*P^*)$
ergibt sich somit:

$$\tilde{\sigma}_e^2 = \frac{1}{n(m-1)} \left\{ \left(\sum_{i,j} y_{ij}^2 - \frac{1}{m} \sum_i y_{i0}^2 \right) + \text{tr} (I - W_0) B \right\},$$

$$\text{wobei } \sum_{i,j} y_{ij}^2 := \sum_{i=1}^{n-1} \sum_{j=1}^m y_{ij}^2 + \sum_{j=1}^{m-1} y_{nj}^2$$

$$y_{i0} := \begin{cases} y_{i.} & \text{für } i < n \\ \sum_{j=1}^{m-1} y_{nj} & \text{für } i = n \end{cases}$$

Berechnung von B:

$$W_0 P Y \otimes W_0 P Y = \frac{1}{m^2} \left[(y_{i0} y_{l0})_{i,l=1,\dots,n} \otimes E_m \right]$$

$$W_0 B W_0 =$$

$$\frac{1}{m} \cdot \left(\begin{array}{c|c} 0_{(n-1)m} & \begin{matrix} c_1 E_m \\ \vdots \\ c_{n-1} E_m \end{matrix} \\ \hline c_1 E_m \dots c_{n-1} E_m & \left(\frac{2(m-1)}{m} c_n + \frac{c_n + c_0}{m} \right) E_m \end{array} \right)$$

$$\text{tr} (I - W_0) B = \text{tr} B - \text{tr} W_0 B W_0 = \frac{m-1}{m} (c_0 - c_n)$$

Die Bedingung (*) führt auf folgendes lineare Gleichungssystem:

$$(i) \quad c_i = \frac{1}{m^2} y_{i.} y_{n0} + \frac{1}{m} c_i \quad \text{für } i < n$$

$$(ii) \quad c_n = \frac{1}{m^2} y_{n0}^2 + \frac{2(m-1)}{m^2} c_n + \frac{c_n + c_0}{m^2} - \frac{1}{m} \tilde{\sigma}_e^2$$

$$(iii) \quad c_n + c_0 = \frac{1}{m^2} y_{n0}^2 + \frac{2(m-1)}{m^2} c_n + \frac{c_n + c_0}{m^2} + \frac{m-1}{m} \tilde{\sigma}_e^2$$

Aus (i) folgt $c_i = \frac{y_i \cdot y_{n0}}{m} \frac{y_{n0}}{m-1}$ für $i < n$,

aus (ii):

$$\begin{aligned} (ii)^* \quad (nm^2 - 2nm) c_n - (n-1) (c_0 - c_n) &= \\ &= ny_{n0}^2 - \frac{m}{m-1} \sum y_{ij}^2 + \frac{m}{m-1} \sum y_{i0}^2 \end{aligned}$$

aus (iii):

$$\begin{aligned} (iii)^* \quad 2nm(m-1) c_n + (nm+n-1) (m-1) (c_0 - c_n) &= \\ &= ny_{n0}^2 + m \cdot \sum y_{ij}^2 - \sum y_{i0}^2 \end{aligned}$$

Auflösen von (ii)* nach c_n und Einsetzen in (iii)* liefert:

$$(nm-n-1) (m-1) (c_0 - c_n) = -ny_{n0}^2 + m \sum y_{ij}^2 - \sum y_{i0}^2$$

Damit ergibt sich für $\hat{\sigma}_e^2$:

$$\hat{\sigma}_e^2 = (nm - n - 1)^{-1} \left\{ \sum y_{ij}^2 - \sum_{i=1}^{n-1} \frac{y_i^2}{m} - \frac{y_{n0}^2}{m-1} \right\}$$

Dieses ist der aus der Varianzanalyse bekannte Schätzer für σ_e^2 .

Wegen $M(PL, P \odot P) = M(PL, P)$ kann man gemäß Corollar zu Lemma 3.2(1) $\tilde{\sigma}_e^2$ auch berechnen, indem man den Beobachtungsvektor PY "ergänzt":

Setze $\tilde{W}_0 := PX_1(X_1'PX_1)^{-1}X_1'P$ orthogonale Projektion auf PL , so ist

$$\begin{aligned} P_{P \perp L}^* Y &= \tilde{W}_0 P^* Y \tilde{W}_0 + (P - \tilde{W}_0) \frac{\langle P - \tilde{W}_0, PYP \rangle}{\langle P - \tilde{W}_0, P - \tilde{W}_0 \rangle} \\ &= \tilde{W}_0 PY \odot \tilde{W}_0 PY + (P - \tilde{W}_0) \frac{((I - \tilde{W}_0)PY, PY)}{\text{tr}(P - \tilde{W}_0)} \end{aligned}$$

Wegen $Rg(PX_1) = Rg(X_1) = n$ ist

$$\text{tr}(P - \tilde{W}_0) = \text{tr} P - Rg(PX_1) = nm - n - 1 \quad \text{und}$$

$\tilde{W}_0 PY = P W_0 (PY + x_0)$ mit $W_0 := X_1(X_1'X_1)^{-1}X_1'$ wie oben und x_0 als eindeutig bestimmte Lösung von

$$(*) \quad x = (I - P)W_0(PY + x), \quad \text{d.h. } x_0 = (0, \dots, 0, \frac{y_{n \odot}}{m-1})'.$$

Für $\tilde{\sigma}_e^2$ ergibt sich hieraus:

$$\begin{aligned} \tilde{\sigma}_e^2 &= (nm - n - 1)^{-1} \cdot \left\{ (PY, PY) - (\tilde{W}_0 PY, PY) \right\} \\ &= (nm - n - 1)^{-1} \cdot \left\{ (PY + x_0, PY + x_0) - (W_0(PY + x_0), PY + x_0) \right\} \\ &= (nm - n - 1)^{-1} \left\{ \sum_{i,j} z_{ij}^2 - \frac{1}{m} \sum_i z_i^2 \right\} \\ &\quad \text{mit } z_{ij} := \begin{cases} y_{ij} & \text{für } (i,j) \neq (n,m) \\ \frac{y_{n \odot}^2}{m-1} & \text{für } (i,j) = (n,m) \end{cases} \\ &= (nm - n - 1)^{-1} \left\{ \sum_{i,j} y_{ij}^2 - \frac{1}{m} \sum_{i=1}^{n-1} y_i^2 - \frac{y_{n \odot}^2}{m-1} \right\} \quad (s.o.) \end{aligned}$$

Es ist also $\tilde{\sigma}_e^2 (Py) = \frac{n(m-1)}{nm-n-1} \hat{\sigma}_e^2 (Py + x_0) .$

Zu beachten ist der Korrekturfaktor $\frac{n(m-1)}{nm-n-1} .$

(vergleiche WILKINSON 1958 b), Seite 363)

3.3.2. Das Modell $y_{ij} = a_i + e_{ij}$

Sei $V^* := S(R^{nm})$, $L^* := [(I_n \otimes E_m), I_{nm}]$ und
 $[\Theta^*] := [\{W \otimes W \mid W \in L^*\}]$

Äquivalente Formulierung:

$$y_{ij} = a_i + e_{ij}, \quad i=1, \dots, n; \quad j=1, \dots, m;$$

$$\text{mit } E_P y_{ij} = 0, \quad \text{Cov}_P(a_1, \dots, a_n)' = \sigma_a^2 \cdot I_n$$

$$\text{Cov}_P(e_{11}, \dots, e_{nm})' = \sigma_e^2 \cdot I_{nm} \quad \text{und} \quad \text{Cov}_P(a, e) = 0.$$

Gesucht sind gleichmäßig optimale Schätzer für σ_a^2 und σ_e^2 .

Wegen $W^2 \in L^* \quad \forall W \in L^*$ ist P_{L^*} GMO zum Modell $M(L^*, \Theta^*)$.

Setze $W_1 := \frac{1}{m} (I_n \otimes E_m)$ und $W_2 := I - W_1$, so ist $W_i W_j = \delta_{ij} W_i$
 und $L^* = [W_1, W_2]$.

$$P_{L^*} Y = W_1 \frac{(W_1 Y, Y)}{\text{tr } W_1} + W_2 \frac{(W_2 Y, Y)}{\text{tr } W_2}$$

$$=: c_1 W_1 + c_2 W_2 = \frac{1}{m} (c_1 - c_2) \cdot (I_n \otimes E_m) + c_2 \cdot I_{nm}.$$

Als Schätzungen für σ_e^2 und σ_a^2 ergeben sich

$$\hat{\sigma}_e^2 = c_2 = \frac{(W_2 Y, Y)}{\text{tr } W_2} = \frac{\sum_{ij} y_{ij}^2 - \frac{1}{m} \sum_i Y_i^2}{n(m-1)}$$

$$\hat{\sigma}_a^2 = \frac{1}{m} (c_1 - c_2) = \frac{\sum_i Y_i^2 - \sum_{ij} y_{ij}^2}{nm(m-1)}$$

Missing-Data-Problem

Sei $P = P' = P^2 = \text{diag} (I_{nm-1}, 0)$ wie in 3.3.1.

Für $W := (I_n \otimes E_m) \in L^*$ ist

$PWPWP = \text{diag} (mE_m, \dots, mE_m, (m-1)E_{m-1}, 0) \notin P^*L^*$, so daß

zum Modell $M(P^*L^*, P^*O^*P^*)$ kein GMO existiert.

Insbesondere existieren keine gleichmäßig besten Schätzer für σ_e^2 und σ_a^2 .

Es lassen sich nur lokal beste Schätzer angeben:

Etwa für $O_0 := [I_{nm}]$ ist $P_{P^*L^*}$ GMO zum Modell $M(P^*L^*, P^*O_0P^*)$.

Zur Berechnung von $P_{P^*L^*}$ läßt sich Theorem 1(7) verwenden:

Wegen $\dim P^*L^* = \dim L^* = 2$ ist $P_{P^*L^*}Y = P^*P_{L^*}(PYP + B)$, wobei $B \in (I - P \otimes P)L^*$ die Gestalt

$B = \text{diag} (0_m, \dots, 0_m, B_m)$ mit

$$B_m = \left(\begin{array}{c|c} 0_{m-1} & \begin{matrix} a \\ \vdots \\ a \end{matrix} \\ \hline a \dots a & a+b \end{array} \right), \quad a, b \in \mathbb{R} \text{ hat.}$$

a und b sind als eindeutig bestimmte Lösungen des linearen Gleichungssystems

(*) $B = (I - P \otimes P)P_{L^*}(PYP + B)$ zu bestimmen.

Es ist $P_{L^*}(PYP + B) = \tilde{c}_1 W_1 + \tilde{c}_2 W_2$ mit

$$\tilde{c}_1 = \frac{1}{n} \left\{ (W_1 P Y, P Y) + \text{tr } W_1 B \right\}$$

$$\tilde{c}_2 = \frac{1}{n(m-1)} \left\{ (W_2 P_Y, P_Y) + \text{tr } W_2 B \right\},$$

$$\text{wobei } \text{tr } W_1 B = \frac{1}{m} \text{tr } E_m B_m = \frac{1}{m} [2(m-1)a + (a+b)]$$

$$\text{und } \text{tr } W_2 B = \text{tr } B - \text{tr } W_1 B = \frac{m-1}{m} [(a+b) - 2a].$$

Die Bedingung (*) ergibt:

$$\begin{aligned} a &= \frac{1}{nm-2} \left\{ (W_1 P_Y, P_Y) - \frac{(W_2 P_Y, P_Y)}{m-1} \right\} \\ &= \frac{1}{(nm-2)(m-1)} \left\{ \sum_i Y_{i\circ}^2 - \sum_{ij}^{\circ} Y_{ij}^2 \right\} \end{aligned}$$

$$\begin{aligned} \text{und } a+b &= \frac{1}{nm-1} \left\{ (W_1 P_Y, P_Y) + (W_2 P_Y, P_Y) \right\} \\ &= \frac{1}{nm-1} \sum_{ij}^{\circ} Y_{ij}^2. \end{aligned}$$

Daraus ergibt sich für $\tilde{\sigma}_e^2$ und $\tilde{\sigma}_a^2$

$$\begin{aligned} \tilde{\sigma}_e^2 = \tilde{c}_2 &= \frac{1}{(nm-2)(m-1)} \left\{ m \sum_{ij}^{\circ} Y_{ij}^2 - \sum_i Y_{i\circ}^2 \right\} \\ &\quad - \frac{1}{(nm-2)(nm-1)} \sum_{ij}^{\circ} Y_{ij}^2 \end{aligned}$$

$$\tilde{\sigma}_a^2 = \frac{1}{m} (\tilde{c}_1 - \tilde{c}_2) = \frac{1}{(nm-2)(m-1)} \left\{ \sum_i Y_{i\circ}^2 - \sum_{ij}^{\circ} Y_{ij}^2 \right\}$$

$\tilde{\sigma}_e^2$ und $\tilde{\sigma}_a^2$ sind erwartungstreue Schätzer für σ_e^2 und σ_a^2 .
Sie sind optimal, falls $\sigma_a^2 = 0$.

Offenbar kann man nicht wie in Beispiel 3.3.1. $\tilde{\sigma}_e^2$ und $\tilde{\sigma}_a^2$ aus $\hat{\sigma}_e^2$ und $\hat{\sigma}_a^2$ berechnen, indem man die bekannten Schätzfunktionen auf einen "erweiterten Beobachtungsvektor" $z = Py + x_0$ anwendet.

$$\text{Es ist } \tilde{\sigma}_e^2(Py) = \frac{nm}{nm-2} \hat{\sigma}_e^2(Py) - \frac{\sum_{ij} y_{ij}^2}{(nm-2)(nm-1)}$$

$$\text{und } \tilde{\sigma}_a^2(Py) = \frac{nm}{nm-2} \hat{\sigma}_a^2(Py) .$$

Berechnung allgemeiner, lokal bester Schätzfunktionen
(Fortsetzung von Beispiel 3.3.2.)

Falls $\text{Cov}_P Y = \sigma_a^2 \cdot (I_n \otimes E_m) + \sigma_e^2 \cdot I_{nm}$ mit $\sigma_e^2 \neq 0$, setze

$c := \sigma_a^2 / \sigma_e^2 \geq 0$, $W := W_\rho := \rho(I_n \otimes E_m) + I_{nm}$ und

$\Theta_\rho := [W \otimes W] \wedge S_+(V^*) \subseteq \Theta^*$.

Wegen $M(L^*, \Theta_\rho) \subseteq M(L^*, \Theta^*)$ ist P_{L^*} GMO zum Modell $M(L^*, \Theta_\rho)$ für alle $\rho \geq 0$.

Falls $\rho \neq 0$, so ist $PWPWP \notin P^*L^*$, also $P_{P^*L^*}$ nicht GMO zum Modell $M(P^*L^*, P^*\Theta_\rho P^*)$. (siehe oben)

Offenbar ist der zuvor behandelte Fall gerade der Spezialfall $\rho=0$.

Falls $\rho \neq 0$, läßt sich der GMO $\tilde{G}$ zum Modell

$M(P^*L^*, P^*\Theta_\rho P^*) \cong M(P^*L^*, P^*(W \otimes W)P^*)$ nicht mit Hilfe von Theorem 1(7) berechnen.

Direkte Berechnung von $\tilde{G} = \tilde{G}_\rho$ für $\rho > 0$:

Es ist $P^*L^* = [P(I_n \otimes E_m)P, P]$ und $\tilde{G} = P_{P^*L^*}^\circ$ Projektion auf L^* entlang $(PWP \otimes PWP)(L^*)^\perp$.

Setze $W_1 := P(I_n \otimes E_m)P = \text{diag}(E_m, \dots, E_m, E_{m-1}, 0)$ und

$W_2 := s \cdot P - t \cdot W_1 \in P^*L^*$, wobei s und $t \in \mathbb{R}$ so gewählt seien, daß $\text{tr}(PWP)^+ W_1 (PWP)^+ W_2 = 0$ ist,

also $t := \text{tr}(PWP)^+ W_1 (PWP)^+$ und $s := \text{tr}(PWP)^+ W_1 (PWP)^+ W_1$.

Dann ist $\tilde{G} P^*Y = \tilde{c}_1 W_1 + \tilde{c}_2 W_2 = (\tilde{c}_1 - t \tilde{c}_2) W_1 + s \tilde{c}_2 W_2$,

so daß $\tilde{\sigma}_e^2 = s \cdot \tilde{c}_2$ und $\tilde{\sigma}_a^2 = \tilde{c}_1 - t \cdot \tilde{c}_2$.

Berechnung von s und t :

$$PWP = \text{diag} (\rho E_m + I_m, \dots, \rho E_m + I_m, \rho E_{m-1} + I_{m-1}, 0)$$

$$(PWP)^+ = \text{diag} (A_m, \dots, A_m, A_{m-1}, 0) \quad \text{mit}$$

$$A_m = (\rho E_m + I_m)^{-1} = -\frac{\rho}{m\rho + 1} E_m + I_m$$

$$(PWP)^+ W_1 (PWP)^+ = \text{diag} (B_m, \dots, B_m, B_{m-1}, 0) \quad \text{mit}$$

$$B_m = \frac{1}{(m\rho + 1)^2} E_m$$

$$(PWP)^+ W_1 (PWP)^+ W_1 = \text{diag} (C_m, \dots, C_m, C_{m-1}, 0) \quad \text{mit}$$

$$C_m = \frac{m}{(m\rho + 1)^2} E_m$$

Daraus ergibt sich:

$$s = \frac{(n-1)m^2}{(m\rho + 1)^2} + \frac{(m-1)^2}{[(m-1)\rho + 1]^2}$$

$$t = \frac{(n-1)m}{(m\rho + 1)^2} + \frac{(m-1)}{[(m-1)\rho + 1]^2}$$

Berechnung von $\tilde{\sigma}_a^2$ und $\tilde{\sigma}_e^2$:

$$\text{Es ist } \tilde{\sigma}_1 = \frac{1}{s} \text{tr} (PWP)^+ W_1 (PWP)^+ PYP \quad \text{und}$$

$$\tilde{\sigma}_2 = [(\text{tr} (PWP)^+ W_2 (PWP)^+ W_2)]^{-1} \text{tr} (PWP)^+ W_2 (PWP)^+ PYP,$$

$$r := \text{tr} (PWP)^+ (PWP)^+$$

$$= \frac{n-1}{(m\rho + 1)^2} + \frac{1}{[(m-1)\rho + 1]^2} + (nm - n - 1)$$

ergibt sich $\text{tr} (PWP)^+ W_2 (PWP)^+ W_2 = s(rs - t^2)$,
so daß

$$\begin{aligned} \tilde{\sigma}_e^2 &= (rs - t^2)^{-1} \left\{ (PWP)^+ W_2 (PWP)^+ P_Y, P_Y \right\} \\ &= (rs - t^2)^{-1} \left\{ s \cdot \left(\sum_{i,j}^{\circ} y_{ij}^2 - \sum_{i=1}^{n-1} \frac{y_{i\cdot}^2}{m} - \frac{y_{n\circ}^2}{m-1} \right) + \right. \\ &\quad \left. + \frac{s - tm}{(m\rho + 1)^2} \sum_{i=1}^{n-1} \frac{y_{i\cdot}^2}{m} + \frac{s - t(m-1)}{[(m-1)\rho + 1]^2} \frac{y_{n\circ}^2}{m-1} \right\} \end{aligned}$$

$$\begin{aligned} \tilde{\sigma}_a^2 &= \frac{1}{s} \cdot \left\{ (PWP)^+ W_1 (PWP)^+ P_Y, P_Y \right\} - \frac{t}{s} \cdot \tilde{\sigma}_e^2 \\ &= (rs - t^2)^{-1} \left\{ -t \left(\sum_{i,j}^{\circ} y_{ij}^2 - \sum_{i=1}^{n-1} \frac{y_{i\cdot}^2}{m} - \frac{y_{n\circ}^2}{m-1} \right) + \right. \\ &\quad \left. + \frac{rm - t}{(m\rho + 1)^2} \sum_{i=1}^{n-1} \frac{y_{i\cdot}^2}{m} + \frac{r(m-1) - t}{[(m-1)\rho + 1]^2} \frac{y_{n\circ}^2}{m-1} \right\} \end{aligned}$$

(vergleiche TOWNSEND & SEARLE 1971, Abschnitt 2.1.)

Für $\rho=0$ ergeben sich die zuvor berechneten Schätzfunktionen.

3.3.3. Das Modell $y_{ij} = \mu + a_i + e_{ij}$

Sei $V^* := S(\mathbb{R}^{nm})$, $L^* := [E_{nm}, (I_n \otimes E_m), I_{nm}]$ und
 $[\Theta^*] := \{W \otimes W \mid W \in L^*\}$

Äquivalente Formulierung:

$$y_{ij} = \mu + a_i + e_{ij}, \quad i=1, \dots, n; \quad j=1, \dots, m;$$

$$\text{mit } E_P y_{ij} = \mu, \quad \text{Cov}_P(a_1, \dots, a_n)' = \sigma_a^2 \cdot I_n$$

$$\text{Cov}_P(e_{11}, \dots, e_{nm})' = \sigma_e^2 \cdot I_{nm} \quad \text{und} \quad \text{Cov}_P(a, e) = 0.$$

Gesucht sind gleichmäßig optimale Schätzer für σ_a^2 und σ_e^2 .

Wegen $W^2 \in L^*$ $\forall W \in L^*$ ist P_{L^*} GMO zum Modell $M(L^*, \Theta^*)$.

Setze $W_0 := \frac{1}{nm} E_{nm}$, $W_1 := \frac{1}{m} (I_n \otimes E_m) - W_0$ und

$$W_2 := I_{nm} - \frac{1}{m} (I_n \otimes E_m) = I_{nm} - W_0 - W_1,$$

so ist $W_i W_j = \delta_{ij} W_i$ für $i, j=1, 2, 3$ und $L^* = [W_0, W_1, W_2]$

$$P_{L^*} Y = \sum_{i=0}^2 c_i W_i \quad \text{mit} \quad c_i := \frac{(W_i Y, Y)}{\text{tr } W_i}.$$

Als Schätzungen für σ_a^2 und σ_e^2 ergeben sich:

$$\hat{c}_e^2 = c_2 = \frac{\sum_{ij} y_{ij}^2 - \frac{1}{m} \sum_i y_i^2}{n(m-1)} \quad (\text{vergl. Beispiel 3.3.2.})$$

$$\hat{\sigma}_a^2 = \frac{c_1 - c_2}{m} = \frac{nm-1}{nm^2(n-1)(m-1)} \sum_i y_{i.}^2 - \frac{1}{nm(n-1)} \sum_{ij} y_{ij}^2 -$$

$$- \frac{1}{nm^2(n-1)} y_{..}^2.$$

$$= \frac{1}{m} \left[\frac{1}{m(n-1)} (\sum_i y_{i.}^2 - \frac{1}{n} y_{..}^2) - \hat{\sigma}_e^2 \right]$$

Missing-Data-Problem

Sei $P = P' = P^2 = \text{diag} (I_{nm-1}, 0)$ wie in 3.3.1. und 3.3.2.

Für $W := (I_n \otimes E_m) \in L^*$ ist

$PWPWP = \text{diag} (mE_m, \dots, mE_m, (m-1)E_{m-1}) \notin P^*L^*$, so daß

zum Modell $M(P^*L^*, P^* \Theta^* P^*)$ kein GMO existiert.

Insbesondere existieren keine gleichmäßig besten Schätzer für σ_e^2 und σ_a^2 . (vergl. Beispiel 3.3.2.)

Es lassen sich nur lokal beste Schätzer angeben:

Für $[e_0] := [I_{nm}]$ ist $P_{P^*L^*}$ GMO zum Modell $M(P^*L^*, P^* \Theta^* P^*)$.

Wegen $\dim P^*L^* = \dim L^* = 3$ ist $P_{P^*L^*} P^*Y = P^*P_{L^*}(PYP + B)$, wobei $B \in (I - P \otimes P)L^*$ die Gestalt

$$B = \left(\begin{array}{cc|c} & & a \\ & & \vdots \\ & & a \\ & 0_{nm-1} & a+b \\ & & \vdots \\ & & a+b \\ \hline a \dots a & a+b \dots a+b & a+b+c \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} m(n-1) \text{ Zeilen} \\ \\ m-1 \text{ Zeilen} \end{array}$$

mit a, b und $c \in \mathbb{R}$.

a, b und c sind als eindeutig bestimmte Lösungen des linearen Gleichungssystems

(*) $B = (I - P \otimes P)P_{L^*}(PYP + B)$ zu bestimmen.

$$\text{Es ist } P_{L*} (PYP + B) = \sum_{i=0}^2 \tilde{c}_i W_i$$

$$\text{mit } \tilde{c}_i = \frac{1}{\text{tr } W_i} \left\{ (W_i P Y, P Y) + \text{tr } W_i B \right\}.$$

Dabei ist

$$\text{tr } W_0 B = \frac{1}{nm} \text{tr } E_{nm} B = \frac{1}{nm} \left\{ 2m(n-1)a + 2(m-1)(a+b) + (a+b+c) \right\}$$

$$\begin{aligned} \text{tr } W_1 B &= \frac{1}{m} \text{tr } (I_n \otimes E_m) B - \frac{1}{nm} \text{tr } E_{nm} B = \\ &= - \frac{2(n-1)}{n} a + \frac{2(n-1)(m-1)}{nm} (a+b) + \frac{n-1}{nm} (a+b+c) \end{aligned}$$

$$\begin{aligned} \text{tr } W_2 B &= \text{tr } B - \frac{1}{m} \text{tr } (I_n \otimes E_m) B = \\ &= - \frac{2(m-1)}{m} (a+b) + \frac{m-1}{m} (a+b+c) \end{aligned}$$

Die Bedingung (*) führt auf das folgende lineare Gleichungssystem:

$$\begin{aligned} \text{(i)} \quad a+b+c &= \frac{1}{nm} \tilde{c}_0 + \frac{n-1}{nm} \tilde{c}_1 + \frac{m-1}{m} \tilde{c}_2 = \\ &= \frac{1}{nm} \left\{ \text{tr } W_0 P Y P + \text{tr } W_1 P Y P + \text{tr } W_2 P Y P \right\} \\ &\quad + \frac{1}{nm} (a+b+c) \end{aligned}$$

$$a+b+c = \frac{1}{nm-1} \sum_{ij} y_{ij}^2 \quad (\text{vergl. Beispiel 3.3.2.})$$

$$\begin{aligned} \text{(ii)} \quad a+b &= \frac{1}{nm} \tilde{c}_0 + \frac{n-1}{nm} \tilde{c}_1 - \frac{1}{m} \tilde{c}_2 = \\ &= \frac{1}{nm} \left\{ \text{tr } W_0 P Y P + \text{tr } W_1 P Y P - \frac{\text{tr } W_2 P Y P}{m-1} \right\} + \\ &\quad + \frac{2}{nm} (a+b) \end{aligned}$$

$$a+b = \frac{1}{(nm-2)(m-1)} \left\{ \sum_i y_{i0}^2 - \sum_{ij} y_{ij}^2 \right\} (\text{vergl. 3.3.2.})$$

$$\begin{aligned}
 \text{(iii)} \quad a &= \frac{1}{nm} \tilde{c}_0 - \frac{1}{nm} \tilde{c}_1 = \\
 &= \frac{1}{nm} \left\{ \text{tr } W_0 \text{PYP} + \frac{\text{tr } W_1 \text{PYP}}{n-1} \right\} + \frac{2}{nm} a \\
 a &= \frac{1}{(nm-2)m(n-1)} \left\{ y_{\ominus\ominus}^2 - \sum_i y_{i\ominus}^2 \right\}
 \end{aligned}$$

Damit ergibt sich für $\tilde{c}_i$:

$$\begin{aligned}
 \tilde{c}_0 &= \frac{1}{nm-2} y_{\ominus\ominus}^2 \\
 \tilde{c}_1 &= \frac{\text{tr } W_1 \text{PYP}}{n-1} - \frac{2}{n} a + \frac{2(m-1)}{nm} (a+b) + \frac{1}{nm} (a+b+c) \\
 &= \frac{1}{(nm-2)(n-1)} \left\{ -y_{\ominus\ominus}^2 + n \sum_i y_{i\ominus}^2 \right\} - \frac{\sum_i y_{ij}^2}{(nm-1)(nm-2)} \\
 \tilde{c}_2 &= \frac{\text{tr } W_2 \text{PYP}}{n(m-1)} - \frac{2}{nm} (a+b) + \frac{1}{nm} (a+b+c) \\
 &= \frac{1}{(nm-2)(m-1)} \left\{ m \sum_i y_{ij}^2 - \sum_i y_{i\ominus}^2 \right\} - \frac{\sum_i y_{ij}^2}{(nm-1)(nm-2)}
 \end{aligned}$$

Es ist

$$\begin{aligned}
 \tilde{\sigma}_e^2 &= \tilde{c}_2 \\
 \tilde{\sigma}_a^2 &= \frac{\tilde{c}_1 - \tilde{c}_2}{m} = \frac{nm-1}{(nm-2)(n-1)(m-1)m} \sum_i y_{i\ominus}^2 \\
 &\quad - \frac{y_{\ominus\ominus}^2}{(nm-2)(n-1)m} - \frac{\sum_i y_{ij}^2}{(nm-2)(m-1)} \\
 &= \frac{1}{m} \left\{ \frac{1}{n-1} \left[\frac{1}{nm-2} \sum_i y_{i\ominus}^2 - \frac{y_{\ominus\ominus}^2}{nm-2} \right] - \tilde{\sigma}_e^2 \right\}
 \end{aligned}$$

Offenbar kann man - ebensowenig wie in Beispiel 3.3.2. -

$\hat{\sigma}_a^2$ und $\hat{\sigma}_e^2$ aus $\hat{\sigma}_e^2$ und $\hat{\sigma}_a^2$ berechnen, indem man die bekannten Schätzfunktionen auf einen "erweiterten Beobachtungsvektor" $z = Py + x_0$ anwendet:

$$\text{Es ist } \tilde{\sigma}_e^2(Py) = \frac{nm}{nm-2} \hat{\sigma}_e^2(Py) - \frac{\sum y_{ij}^2}{(nm-2)(nm-1)}$$

$$\text{und } \tilde{\sigma}_a^2(Py) = \frac{nm}{nm-2} \hat{\sigma}_a^2(Py) \quad (\text{vergleiche Beispiel 3.3.2.})$$

Berechnung lokal bester Schätzfunktionen

(Fortsetzung von Beispiel 3.3.3.)

Falls $\text{Cov}_{PY} = \sigma_a^2 (I_n \otimes E_m) + \sigma_e^2 \cdot I_{nm}$ mit $\sigma_e^2 > 0$,

setze $\rho := \sigma_a^2 / \sigma_e^2 \geq 0$, $W := W_\rho := \rho (I_n \otimes E_m) + I_{nm}$ und

$[\Theta_\rho] := [W \otimes W] \subseteq [\Theta^*]$ (wie in 3.3.2.)

Für $\rho \neq 0$ ist $PWPWP \notin \mathbb{P}^*L^*$, also $P_{\mathbb{P}^*L^*}$ nicht GMO zum Modell $M(\mathbb{P}^*L^*, \mathbb{P}^*C_\rho \mathbb{P}^*)$.

Direkte Berechnung von $G = G_\rho$ für $\rho > 0$.

(Es werden erwartungstreue Schätzer für σ_e^2 und σ_a^2 berechnet. Sie sind optimal, falls $\rho = \sigma_a^2 / \sigma_e^2$ und $\mu = 0$ die wahren Parameterwerte sind.)

Für $\rho = 0$ ergibt sich gerade der zuvor behandelte Spezialfall.

Setze $W_1 := P(I_n \otimes E_m)P$

$W_2 := sP - tW_1$ mit s und t wie in 3.3.2.

$W_0 := PE_{nm}P - W_1$

so ist $\langle W_1, W_2 \rangle_0 := \text{tr}(W_1 (PWP)^+ W_2 (PWP)^+) = 0$.

Wegen $(PWP)^+ (PEP) (PWP)^+ =$

$$\begin{pmatrix} A & . & . & A & B & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ A & . & . & A & B & . \\ B' & . & . & B' & C & . \\ 0 & . & . & . & . & 0 \end{pmatrix}$$

mit $A = (I_m - \frac{c}{m\rho+1} E_m) E_m (I_m - \frac{c}{m\rho+1} E_m) = \frac{1}{(m\rho+1)^2} E_m$

$$B = (I_m - \frac{\rho}{m\rho+1} E_m) E_m \begin{pmatrix} I_{m-1} - \frac{\rho}{(m-1)\rho+1} E_{m-1} & 0 \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

$$= \frac{1}{(m\rho+1) [(m-1)\rho+1]} \begin{pmatrix} E_{m-1} & 0 \\ 0 & \dots & 0 \end{pmatrix}$$

$$C = (I_{m-1} - \frac{\rho}{(m-1)\rho+1} E_{m-1}) E_{m-1} (I_{m-1} - \frac{\rho}{(m-1)\rho+1} E_{m-1})$$

$$= \frac{1}{[(m-1)\rho+1]^2} E_{m-1}$$

$$\text{ist } \langle \text{PEP}, W_1 \rangle_\rho = n \cdot \text{tr } A E_m + \text{tr } C E_{m-1} =$$

$$= \frac{nm^2}{(m\rho+1)^2} + \frac{(m-1)^2}{[(m-1)\rho+1]^2} = s = \langle W_1, W_1 \rangle_c,$$

$$\text{und } \langle \text{PEP}, P \rangle_c = n \cdot \text{tr } A + \text{tr } C =$$

$$= \frac{nm}{(m\rho+1)^2} + \frac{m-1}{[(m-1)\rho+1]^2} = t = \langle W_1, P \rangle_c,$$

$$\text{so da\ss } \langle W_1, W_0 \rangle_c = \langle W_1, \text{PEP} \rangle_\rho - \langle W_1, W_1 \rangle_c = 0$$

$$\text{und } \langle W_2, W_0 \rangle_c = \langle W_2, \text{PEP} \rangle_\rho - \langle W_2, W_1 \rangle_\rho$$

$$= s \langle P, \text{PEP} \rangle_c - t \langle W_1, \text{PEP} \rangle_\rho = st - ts = 0.$$

$$\text{Damit ist } G_0 \text{PYP} = \sum_{i=0}^2 \tilde{c}_i \cdot W_i$$

$$= \tilde{c}_0 \cdot \text{PEP} + (s \cdot \tilde{c}_2 + \tilde{c}_1 - \tilde{c}_0) W_1 - t \cdot \tilde{c}_2 P,$$

$$\text{mit } \tilde{c}_i = \frac{\langle W_i, \text{PYP} \rangle_\rho}{\langle W_i, W_i \rangle_c}.$$

Dabei sind $\tilde{c}_1$ und $\tilde{c}_2$ wie in Beispiel 3.3.2.,

$$\tilde{c}_e^2 = t \cdot \tilde{c}_2 \quad (\text{wie in 3.3.2.}) \quad \text{und} \quad \tilde{c}_a^2 = s \cdot \tilde{c}_2 + \tilde{c}_1 - \tilde{c}_0.$$

Berechnung von $\tilde{C}_0$:

$$\langle W_0, W_0 \rangle_\rho = \langle PEP, W_0 \rangle_\rho = \langle PEP, PEP \rangle_\rho - \langle PEP, W_1 \rangle_\rho =: q - s$$

mit s wie oben

$$\text{und } q = \text{tr } (PEP) (PWP)^+ (PEP) (PWP)^+$$

$$\begin{aligned} &= (n-1)^2 \cdot \text{tr } AE_m + (n-1) \cdot \text{tr } BPE_m + (n-1) \cdot \text{tr } B'E_m P \\ &\quad + \text{tr } CE_{m-1} \\ &= \frac{(n-1)^2 m^2}{(m_0+1)^2} + 2 \frac{m(n-1)(m-1)}{(m_0+1) [(m-1)_0+1]} + \frac{(m-1)^2}{[(m-1)_0+1]^2} \\ &= \left[\frac{m(n-1)}{m_0+1} + \frac{m-1}{(m-1)_0+1} \right]^2. \end{aligned}$$

$$(PWP)^+ (PEP) (PWP)^+ P_Y = (\underbrace{z_1, \dots, z_1}_{m(n-1)}, \underbrace{z_2, \dots, z_2}_{m-1}, 0)' \quad \text{mit}$$

$$z_1 = \frac{1}{(m_0+1)^2} \sum_{i=1}^{n-1} Y_{i\cdot} + \frac{1}{(m_0+1) [(m-1)_0+1]^2} Y_{n\cdot}$$

$$z_2 = \frac{1}{(m_0+1) [(m-1)_0+1]} \sum_{i=1}^{n-1} Y_{i\cdot} + \frac{1}{[(m-1)_0+1]^2} Y_{n\cdot}$$

$$\text{tr } (PWP)^+ (PEP) (PWP)^+ P_Y P_Y = \left((PWP)^+ (PEP) (PWP)^+ P_Y, P_Y \right)$$

$$= \frac{\left(\sum_{i=1}^{n-1} Y_{i\cdot} \right)^2}{(m_0+1)^2} + 2 \frac{\left(\sum_{i=1}^{n-1} Y_{i\cdot} \right) \cdot Y_{n\cdot}}{(m_0+1) [(m-1)_0+1]} + \frac{Y_{n\cdot}^2}{[(m-1)_0+1]^2}$$

$$= \left[\sum_{i=1}^{n-1} \frac{Y_{i\cdot}}{m_0+1} + \frac{Y_{n\cdot}}{(m-1)_0+1} \right]^2$$

Damit ergibt sich für $\tilde{c}_0$ und $\tilde{\sigma}_a^2$:

$$\tilde{c}_0 = \frac{1}{q-s} \left\{ \left(\sum_{i=1}^{n-1} \frac{y_{i.}}{m\rho+1} + \frac{y_{n0}}{(m-1)\rho+1} \right)^2 - \frac{\sum_{i=1}^{n-1} y_{i.}^2}{(m\rho+1)^2} - \frac{y_{n0}^2}{[(m-1)\rho+1]^2} \right\}$$

$$\begin{aligned} \tilde{\sigma}_a^2 = & \frac{-t}{rs-t^2} \left(\sum_{i,j} y_{ij}^2 - \sum_{i=1}^{n-1} \frac{y_{i.}^2}{m} - \frac{y_{n0}^2}{m-1} \right) \\ & - \frac{1}{q-s} \left(\sum_{i=1}^{n-1} \frac{y_{i.}}{m\rho+1} + \frac{y_{n0}}{(m-1)\rho+1} \right)^2 \\ & + \sum_{i=1}^{n-1} \frac{y_{i.}^2}{m} \cdot \left(\frac{sm-t}{rs-t^2} + \frac{m}{q-s} \right) \cdot \frac{1}{(m\rho+1)^2} \\ & + \frac{y_{n0}^2}{m-1} \cdot \left(\frac{s(m-1)-t}{rs-t^2} + \frac{m-1}{q-s} \right) \cdot \frac{1}{[(m-1)\rho+1]^2} \end{aligned}$$

Für $\rho=0$ ergibt sich gerade der zuvor berechnete Wert.

(vergleiche LAMOTTE 1973, (32) und (33))

$$\text{mit } x_1 = \frac{s}{rs-t^2} + \frac{1}{q-s}, \quad x_2 = \frac{-t}{rs-t^2} \quad \text{und } c_0=0 \quad (\mu=0)$$

$$\tilde{\sigma}_a^2 = x_1 Y' M_1 Y + x_2 Y' M_2 Y$$

Anmerkungen zu 3.3.

- 1) Für lineare Modelle der Form 3.3.0. mit $\mathcal{H}(X_2) \subseteq \mathcal{H}(X_1)$ (vergleiche DRYGAS 1977, Beispiel A: "RCR-Modell") ist auch im Fall unvollständiger Daten die Fehlervarianz σ_e^2 gleichmäßig optimal schätzbar. Wegen Theorem 3(3)iii) ist der BQUE $\tilde{\sigma}_e^2$ invariant gegen Translationen von $E_P y$ in $\mathcal{H}(X_1)$.
Insbesondere läßt sich $\tilde{\sigma}_e^2$ berechnen, indem man zunächst die fehlenden Meßwerte "durch Schätzwerte ersetzt". Zu berücksichtigen ist ein Korrekturfaktor $Rg(X_1) / (Rg(X_1) - Rg(I - P))$, der der tatsächlichen Zunahme der Varianz Rechnung trägt.

Dieses aus der Varianzanalyse bekannte Ergebnis (siehe z.B. WILKINSON 1958 b), Seite 363) wurde hier aus der GM-Theorie für lineare Schätzer hergeleitet.
- 2) Die Beispiele 3.3.2. und 3.3.3. weisen auf eine neue Möglichkeit hin, quadratische Schätzfunktionen für den Fall unvollständiger Daten zu ermitteln.
Es werden dabei nicht - wie bei der Berechnung linearer Schätzer (siehe Beispiele 1.3.) - die Schätzwerte unmittelbar mit Hilfe eines aufgefüllten Datensatzes berechnet, sondern Schätzfunktionen entwickelt, die für bestimmte Parameterwerte optimale Schätzer liefern.

Die Beispiele sind bewußt als Spezialfälle bereits bekannter (und gelöster) Probleme gewählt worden (siehe TOWNSEND & SEARLE 1971 sowie LAMOTTE 1973 und SWALLOW & SEARLE 1978), um die Zuverlässigkeit des methodischen Ansatzes zu demonstrieren. Sie reichen nicht aus, um über den numerischen Vorteil dieses Verfahrens verlässliche Aussagen machen zu können, zumal die Komplexität mit der Zahl der fehlenden Werte vermutlich schnell ansteigt.

4. Das allgemeine multivariate Regressionsmodell

4.0. Das Modell

Sei $V := \mathbb{R}^{np}$, $L := \mathcal{M}(I_p \otimes X)$, X bekannte reelle $n \times q$ -Matrix,
 $\Theta := \{\Sigma \otimes Q \mid \Sigma \in S_+(\mathbb{R}^p)\}$, $Q \in S_+(\mathbb{R}^n)$ bekannt.

Äquivalente Formulierung:

$Y := (y_{(1)}, \dots, y_{(p)})$ Matrix der Beobachtungswerte mit $y_{(i)} \in \mathbb{R}^n$,

$E_P Y_{(i)} = X \beta_{(i)}$ mit $\beta_{(i)} = \beta_{(i)}(P) \in \mathbb{R}^q$ und

$\text{Cov}_P Y_{(i)} = Q \quad \forall i=1, \dots, p, \forall P \in \mathcal{P}$.

$Y := (y'_{(1)}, \dots, y'_{(p)})'$.

4.1. GM-Schätzer für $E_P Y$ und Σ

Sei $L_0 := \mathcal{M}(X) \subseteq \mathbb{R}^n$,

W_0 der Projektor auf L_0 entlang $QL_0^\perp \oplus (L_0^\perp \cap K(Q))$,

insbesondere ist also W_0 GMO zum Modell $M(L_0, Q)$ in $\mathbb{R}^n$.

Behauptung: $G := I_p \otimes W_0$ ist GMO zum Modell $M(L, \Theta)$ in V .

Beweis: $G: V \rightarrow L$ idempotent und

$G(\Sigma \otimes Q) = \Sigma \otimes W_0 Q$ symmetrisch $\forall \Sigma \in S_+(\mathbb{R}^p)$.

Mit Lemma 0(4) (Symmetriekriterium) folgt die Behauptung.

Bemerkung:

Falls man ein allgemeineres Modell betrachtet,
etwa $M(L, \theta)$ in V mit V und L wie oben,

$\theta := \{\Sigma \otimes Q \mid \Sigma \in S_+(\mathbb{R}^p), Q \in S\}$, wobei $S \subseteq S_+(\mathbb{R}^n)$ eine beliebige
Teilmenge sei, so ist gemäß Kapitel 2. S durch ein $\bar{Q} \in S_+(\mathbb{R}^n)$
zu ersetzen, für das $\bar{Q}\mathbb{R}^n = \sum Q\mathbb{R}^n$ gelten muß.

Folgende Aussagen sind dann paarweise äquivalent:

- i) $\exists Q \in S_+(\mathbb{R}^n): M(L_0, S) \cong M(L_0, \bar{Q})$ in $\mathbb{R}^n$
- ii) W_0 ist GMO zu $M(L_0, S)$ in $\mathbb{R}^n$ (W_0 GMO zu $M(L_0, \bar{Q})$)
- iii) es existiert ein GMO zu $M(L_0, S)$ in $\mathbb{R}^n$
- iv) $\exists \bar{Q} \in S_+(\mathbb{R}^n): M(L, \theta) \cong M(L, I_p \otimes \bar{Q})$
- v) $G := I_p \otimes W_0$ ist GMO zu $M(L, \theta)$ in $V = \mathbb{R}^{np}$
- vi) es existiert ein GMO zu $M(L, \theta)$ in V .

Sei $V^* := S(V)$, $L^* := L_1 + L_2$ mit

$$L_1 := S(L) = \{(I_p \otimes X) \wedge (I_p \otimes X') \mid \Lambda \in S(\mathbb{R}^n)\},$$

$$L_2 := [\theta] = [\{W_{ij} \mid i, j=1, \dots, p, i \leq j\}],$$

$$- \text{wobei } W_{ij} := \frac{1}{2}(e_i e_j' + e_j e_i') \otimes Q = W_{ij}' = W_{ji} \geq 0 -$$

und $\theta^* := [\{W \otimes W \mid W \in L^*\}] \cap S_+(V^*)$ - (siehe Abschnitt 3.1.1.) -

Sei $T := X \Lambda X' + Q \in S_+(\mathbb{R}^n)$ so daß $T\mathbb{R}^n = L_0 + Q\mathbb{R}^n$, so ist

$$\bar{W} := I_p \otimes T \in L^* \cap S_+(V) \text{ und } \bar{W}V = \sum_{W \in L^*} WV,$$

d.h. $\bar{W}$ erfüllt die Voraussetzungen von Lemma 3.1(3).

Wegen Theorem 3(1) existiert ein GMO zum Modell $M(L^*, \theta^*)$

$$\iff A\bar{W}^+A \in L^* \quad \forall A \in L^*.$$

Behauptung: Für $\bar{W} := I_p \otimes T = I_p \otimes (X\Lambda X' + Q)$ mit $\Lambda = \Lambda' > 0$ gilt:

$$A\bar{W}^+A \in L^* \quad \forall A \in L^*$$

Beweis: siehe DRYGAS 1977, Beispiel B

Berechnung von $\hat{\Sigma}$

Sei W_0 der Projektor auf L_0 entlang $QL_0^\perp \oplus (L_0^\perp \cap K(Q))$ wie oben, mit Corollar zu Lemma 3.1(5) ist dann

$$P_{L_1}^\circ = (I_p \otimes W_0) \otimes (I_p \otimes W_0') \text{ GMO zu } M(L_1, \bar{W} \otimes \bar{W}) \cong M(L_1, \Theta^*) \text{ in } V^*.$$

Da W_0Q symmetrisch, also $(I_p \otimes W_0)(\Sigma \otimes Q) = \Sigma \otimes W_0Q$ symmetrisch $\forall \Sigma = \Sigma'$, gilt:

$$\begin{aligned} \bar{W}_{ij} &:= (I - P_{L_1}^\circ) \bar{W}_{ij} = \frac{1}{2}(e_i e_j' + e_j e_i') \otimes (Q - W_0 Q W_0') = \\ &= \frac{1}{2}(e_i e_j' + e_j e_i') \otimes (I - W_0)Q(I - W_0)' \geq 0. \end{aligned}$$

Wegen $\text{tr } \bar{W}_{ij} \bar{W}_{kl}^+ \bar{W}_{kl}^+ = 0$ für $(i,j) \neq (k,l)$

folgt mit Theorem 3(2) und Lemma 3.1(9) für $\bar{W} := I_p \otimes Q$:

$$\begin{aligned} \hat{\sigma}_{ij} &= \frac{(\bar{W}^+ \bar{W}_{ij} \bar{W}^+ (I - W_0)y, (I - W_0)y)}{\text{tr } \bar{W}^+ \bar{W}_{ij} \bar{W}^+ \bar{W}_{ij}} \\ &= \frac{(Q^+ (I - W_0)y, (I - W_0)y)}{\text{tr } (I - W_0)' Q^+} \quad \text{für } i, j = 1, \dots, p, \end{aligned}$$

so daß für $\hat{\Sigma} = (\hat{\sigma}_{ij})_{i,j=1,\dots,p}$ gilt:

$$\hat{\Sigma} = \frac{1}{\text{tr } (I - W_0)' Q^+} Y' (I - W_0)' Q^+ (I - W_0) Y.$$

(vergleiche DRYGAS 1977, Beispiel B)

4.2. Missing-Data-Problem

Statt $y = (y'_{(1)}, \dots, y'_{(p)})'$ werde nur

$P_y := ((P_1 y_{(1)})', \dots, (P_p y_{(p)})')'$ beobachtet mit $P_i = P_i^1 = P_i^2$,

$P_i = \text{diag}(\Delta_{1i}, \dots, \Delta_{ni})$ ($i=1, \dots, p$) - (siehe Abschnitt 1.2.0.)

Es ist $E_p P_y = \text{diag}(P_1 X, \dots, P_p X) \beta$

und $\text{Cov}_p P_y = P(\Sigma \otimes Q)P = \sum \sigma_{ij} (e_i e_j' \otimes P_i Q P_j)$.

Sei W_i der Projektor auf $L_i := P_i L_0$ entlang

$(P_i Q P_i) L_i^\perp \oplus (L_i^\perp \cap K(P_i Q P_i))$ ($i=1, \dots, p$),

d.h. W_i ist GMO zu $M(L_i, P_i Q P_i)$ in $\mathbb{R}^n$.

Behauptung: $\tilde{G} := \text{diag}(W_1, \dots, W_p)$ ist GMO
zum Modell $M(PL, P(I_p \otimes Q)P)$ in $V = \mathbb{R}^{np}$

Beweis: $\tilde{G}: V \rightarrow PL$ idempotent und

$\tilde{G}(I \otimes Q)P = \text{diag}(W_1 P_1 Q P_1, \dots, W_p P_p Q P_p)$
symmetrisch, da nach Wahl der W_i $W_i P_i Q P_i$ symm. $\forall i$.

(vergleiche DRYGAS 1976)

Notwendig und hinreichend für die Existenz eines GMO zum
Modell $M(PL, P \otimes P)$ in V ist:

$\tilde{G}P(\Sigma \otimes Q)P$ symmetrisch $\forall \Sigma = \Sigma' \geq 0$, bzw. äquivalent:

$$(\Delta) \quad W_i P_i Q P_j = P_i Q P_j W_j' \quad \forall i, j=1, \dots, p.$$

Bemerkung 1 Falls $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$, so ist

$\tilde{G}P(\Sigma \otimes Q)P = \sum \sigma_{ii} (e_i e_i' \otimes W_i P_i Q P_i)$ symmetrisch,

d.h. $\tilde{G}P_y = ((W_1 P_1 y_{(1)})', \dots, (W_p P_p y_{(p)})')'$
ist lokal optimal für diagonales Σ .

Bemerkung 2 Falls $Q = I$, so ist $W_i = W'_i$ ($i=1, \dots, p$) und
 (Δ) ist äquivalent zu $P_i W_i = W_j P_i \quad \forall i, j=1, \dots, p$.

Folgende Aussagen sind dann paarweise äquivalent:

- i) $P_i W_j = W_i P_j \quad \forall i, j=1, \dots, p$
- ii) $P_i W_j = W_i W_j \quad \forall i, j=1, \dots, p$
- iii) $P_i L_j \subseteq L_i \quad \forall i, j=1, \dots, p$.

(iii) ist die Charakterisierung von DRYGAS 1976)

Bemerkung 3 Die Bedingung (Δ) ist z.B. erfüllt, falls
 $P_1 = \dots = P_p = P$, d.h. ganze Zeilen von Y
 fehlen,

(siehe auch McDonald 1971, Theorem 1)

oder falls $P_i Q P_j = \delta_{ij} P_i Q P_i \quad \forall i, j=1, \dots, p$.

I.a. ist (Δ) vermutlich kaum je erfüllt, so daß
 im Falle unvollständiger Daten keine gleichmäßig
 besten Schätzer für $E_p P y$ existieren.

Wegen Theorem 3(3) iii) sind dann die (lokal)
 besten Schätzer für Σ bzw. σ_{ij} ($i, j=1, \dots, p$) nicht
 invariant gegen Translationen von $E_p P y$.

"Ersetzen fehlender Daten"

Falls $M(PL, P(I_p \otimes Q)P)$ die Voraussetzungen von Theorem 1(8)
 erfüllt, also insbesondere " $\dim PL = \dim L$ ", so kann man $\hat{G}$
 mit Hilfe von $G = I_p \otimes W_0$ berechnen.

Da $\hat{G} = \text{diag}(W_1, \dots, W_p)$, das multivariate Modell $M(PL, P(I_p \otimes Q)P)$
 also in p univariate Modelle $M(L_i, P_i Q P_i)$ zerfällt, geht es nur
 um die Berechnung der W_i ($i=1, \dots, p$).

Falls für $i=1, \dots, p$ das Modell $M(L_i, P_i Q P_i)$ die Voraussetzungen
 von Theorem 1(8) erfüllt, so ist

$W_i P_i y_{(i)} = P_i W_0 (P_i y_{(i)} + x_{(i)})$, wobei $x_{(i)}$ Lösung von

$x = (I - P_i) W_0 (P_i y_{(i)} + x)$ ist.

Zum Ersetzen der fehlenden Werte $(I-P_i)Y_{(i)}$ aus Spalte i der Datenmatrix Y werden also nur die Beobachtungen $P_i Y_{(i)}$ aus derselben Spalte herangezogen.

Beste quadratische Schätzer für Σ

Wegen Theorem 3(1) existiert ein GMO zum Modell $M(P^*L^*, P^*\Theta^*P^*)$

$$\Longleftrightarrow A(P\bar{W}P)^+ A \in P^*L^* \quad \forall A \in P^*L^*,$$

wobei $\bar{W} := I_p \otimes T$ wie oben gewählt sei.

$A \in P^*L^*$ hat die Form

$$A = \text{diag}(P_1 X, \dots, P_p X) \wedge \text{diag}(X' P_1, \dots, X' P_p) + \sum \sigma_{ij} e_i e_j' \otimes P_i Q P_j$$

und es ist $(P\bar{W}P)^+ = \text{diag}((P_1 T P_1)^+, \dots, (P_p T P_p)^+)$, so daß

$$A(P\bar{W}P)^+ A =$$

$$(I) \quad P(I_p \otimes X) [\wedge \text{diag}(X' (P_1 T P_1)^+ X, \dots, X' (P_p T P_p)^+ X) \wedge] (I_p \otimes X) P$$

$$(II) \quad \left\{ \begin{array}{l} + \text{diag}(P_1 X, \dots, P_p X) \wedge (\sum \sigma_{ij} e_i e_j' \otimes X' (P_i T P_i)^+ P_i Q P_j) \\ + (\sum \sigma_{ij} e_i e_j' \otimes P_i Q P_j (P_j T P_j)^+ X) \wedge \text{diag}(X' P_1, \dots, X' P_p) \end{array} \right.$$

$$(III) + \sum_{ij} e_i e_j' \otimes (\sum_k \sigma_{ik} \sigma_{jk} (P_i Q P_k) (P_k T P_k)^+ (P_k Q P_j)).$$

Offensichtlich ist für beliebiges $\Sigma = \Sigma' = (\sigma_{ij})_{i,j=1,\dots,p}$

$$(I) \in P^*L_1 \subseteq P^*L^*.$$

Behauptung: Falls $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ mit $\sigma_{ii} \geq 0$ beliebig, so ist (II) und (III) $\in P^*L^*$.

Beweis: Es werden folgende Umformungen benutzt (vergleiche DRYGAS 1977, (7.8) und (7.9)):

$$X'(P_i TP_i)^+ (P_i QP_i) = X'P_i - X'(P_i TP_i)^+ P_i XX'P_i$$

und

$$\begin{aligned} (P_i QP_i) (P_i TP_i)^+ (P_i QP_i) \\ = P_i QP_i - P_i XX'P_i - P_i XX'P_i (P_i TP_i)^+ P_i XX'P_i. \end{aligned}$$

(II) hat die Form

$$\begin{aligned} P(I_p \otimes X) \Lambda \left[\sum \sigma_{ii} e_i e_i' \otimes I_n \right] (I_p \otimes X'P_i) - \\ - \sum \sigma_{ii} e_i e_i' \otimes X'(P_i TP_i)^+ P_i X (I_p \otimes X'P_i) \Lambda \\ - \sum \sigma_{ii} (I_p \otimes P_i X) (e_i e_i' \otimes I_n) - \\ - \sum \sigma_{ii} (I_p \otimes P_i X) (e_i e_i' \otimes X'P_i (P_i TP_i)^+ P_i X) \Lambda (I_p \otimes X) P \\ = P(I_p \otimes X) \left[\Lambda (\Sigma \otimes I_n) - (\Sigma \otimes I_n) \Lambda \right] (I_p \otimes X') P \\ - P(I_p \otimes X) \left[\Lambda (\Sigma \otimes I_n) \text{diag}(X'(P_i TP_i)^+ X) - \right. \\ \left. - \text{diag}(X'(P_i TP_i)^+ X) (\Sigma \otimes I_n) \Lambda \right] (I_p \otimes X') P, \end{aligned}$$

so daß (II) $\in P^*L_1 \subseteq P^*L^*$.

(III) hat die Form

$$\begin{aligned} \sum \sigma_{ii} e_i e_i' \otimes P_i QP_i - \sum \sigma_{ii} (e_i e_i' \otimes P_i X) (I_p \otimes X'P_i) \\ + \sum \sigma_{ii} (e_i e_i' \otimes P_i X) (I_p \otimes X' (P_i TP_i)^+ X) (I_p \otimes X'P_i) \\ = P(\Sigma^2 \otimes Q) P - P(I_p \otimes X) (\Sigma \otimes I_p) (I_p \otimes X') P + \\ + P(I_p \otimes X) \left[\sum \sigma_{ii} (e_i e_i' \otimes X'(P_i TP_i)^+ X) \right] (I_p \otimes X') P, \end{aligned}$$

so daß (III) $\in P^*L^*$.

Falls man also ein Modell $M(P^*L^*, P^*0^*P^*)$ mit $L^* := L_1 + [0_0]$ und $0_0 := \{\Sigma \otimes Q \mid \Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp}) \geq 0\}$ betrachtet, so existiert zu beliebigem $P = P' = P^2$ ein GMO ($(III) \in P^*L^*$!).

D.h. es existieren $\forall (\sigma_{11}, \dots, \sigma_{pp})$ gleichmäßig beste quadratische erwartungstreue Schätzer für $\sigma_{11}, \dots, \sigma_{pp}$.

Wegen Theorem 3(3) sind diese Schätzer invariant gegen Translationen von $E_P Py$, insbesondere hängen sie nicht von $\beta = (\beta'_{(1)}, \dots, \beta'_{(p)})'$ ab.

I.a. sind jedoch (II) und (III) $\notin P^*L^*$, so daß keine ($\forall \beta \in \mathbb{R}^{pq}$, $\Sigma = \Sigma' \geq 0$) gleichmäßig besten Schätzer für $\Sigma = (\sigma_{ij})$ bzw. für σ_{ij} ($i, j=1, \dots, p$) existieren.

Die Anwendung von Theorem 3(2) (ii) auf

$$\tilde{W}_{ij} := \frac{1}{2}(e_i e_j' + e_j e_i') \otimes (P_i Q P_j - W_i Q W_j') \quad (W_1, \dots, W_p \text{ wie oben})$$

und $\bar{W} := \text{diag}(P_1 T_1 P_1, \dots, P_p T_p P_p)$ mit $T_i := X \beta_{(i)} \beta_{(i)}' X' + Q$ ($\beta_{(i)} \in \mathbb{R}^q$ fest gewählt für $i=1, \dots, p$)

liefert die bei $E_{P^*(i)} y = X \beta_{(i)}$ ($i=1, \dots, p$), $\sigma_{ij} = \delta_{ij} \sigma_{ii} \geq 0$ beliebig lokal besten Schätzer für σ_{ij} bzw. für Σ :

$$\tilde{c}_{ij} = \frac{\frac{1}{2} (P_i Y_{(i)})' (P_i T_i P_i)^+ [P_i Q P_j - W_i Q W_j'] (P_j T_j P_j)^+ P_j Y_{(j)}}{\text{tr } \bar{W}^+ \tilde{W}_{ij} \bar{W}^+ \tilde{W}_{ij}}$$

Die bei $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ ($\sigma_{ii} \geq 0$ beliebig) lokal besten invarianten Schätzer für σ_{ij} erhält man durch Anwendung von Theorem 3(3) bzw. Lemma 3.1(10) auf

$$\tilde{W}_{ij} := \frac{1}{2} (e_i e_j' + e_j e_i') \otimes (P_i - W_i) Q (P_j - W_j)' \quad \text{und}$$

$$\bar{W} := \text{diag} (P_1 Q P_1, \dots, P_p Q P_p):$$

$$\tilde{\sigma}_{ij} := \frac{\frac{1}{2} [(I - W_i) P_i Y_{(i)}]' (P_i Q P_i)^+ Q (P_j Q P_j)^+ [(I - W_j) P_j Y_{(j)}]}{\text{tr } \bar{W}^+ \tilde{W}_{ij} \bar{W}^+ \tilde{W}_{ij}}.$$

Wegen Theorem 3(3) ist $\tilde{\sigma}_{ij} = \tilde{\sigma}_{ij} \quad \forall i, j = 1, \dots, p \quad \forall Y \in V$ genau dann wenn die Bedingung (Δ) erfüllt ist.

Für beliebiges $Y \in V$, $P = P' = P^2$ ist jedoch $\tilde{\sigma}_{ii} = \tilde{\sigma}_{ii} \quad (i=1, \dots, p)$.

"Ersetzen fehlender Daten"

Falls $Q=I$, $\text{Rg}(P_i X) = \text{Rg}(X)$ für $i=1, \dots, p$, so folgt aus Lemma 3.2(1) mit $\bar{W} = \bar{W}^+ = P$

wegen $\text{tr } W_{ij} (P_j - W_j) W_{kl} (P_l - W_l) = 0$ für $(i, j) \neq (k, l)$:

$$\begin{aligned} \tilde{\sigma}_{ij} &= \frac{\frac{1}{2} [(I - W_i) P_i Y_{(i)}]' [(I - W_j) P_j Y_{(j)}]}{\text{tr } \frac{1}{2} [e_i e_j' \otimes (P_i - W_i) (P_j - W_j) + e_j e_i' \otimes (P_j - W_j) (P_i - W_i)]} \\ &= \frac{(P_i Y_{(i)} + x_{(i)})' (I - W_i) (P_j Y_{(j)} + x_{(j)})}{\text{tr } (P_i - W_i) (P_j - W_j)} \end{aligned}$$

wobei $x_{(i)}$ ($i=1,..p$) jeweils die eindeutig bestimmte Lösung
von $x = (I-P_i)W_o(P_i Y_{(i)} + x)$ ist.

4.3. Beispiel

Sei $X := (1, \dots, 1)'$, $Q := I_n$,

d.h. $E_p Y_{ik} = \mu_i \in \mathbb{R} \quad (k=1, \dots, n, i=1, \dots, p)$

und $\text{Cov}_p Y_{(i)} = I_n \quad (i=1, \dots, p)$.

Es ist $W_0 = \frac{1}{n} E_n = W_0'$ die orthogonale Projektion auf $L_0 = \mathcal{H}(X)$,
sodaß im Fall vollständiger Daten

$$\hat{\mu}_i = \frac{1}{n} \sum_{k=1}^n Y_{ik} =: \frac{1}{n} Y_{\cdot i}$$

$\forall \mu = (\mu_1, \dots, \mu_p)' \in \mathbb{R}^p$, $\forall \Sigma = \Sigma' \geq 0$ gleichmäßig bester linearer
Schätzer für $\mu_i \quad (i=1, \dots, p)$ ist (siehe Abschnitt 4.1.)

$$\text{und } \hat{\Sigma} = \frac{1}{\text{tr}(I - W_0)} Y' (I - W_0) Y = \frac{1}{n-1} Y' (I_n - \frac{1}{n} E_n) Y$$

$\forall \mu \in \mathbb{R}^p$, $\forall \Sigma = \Sigma' \geq 0$ gleichmäßig bester quadratischer Schätzer
für Σ ist,

$$\begin{aligned} \text{bzw. } \hat{\sigma}_{ij} &= \frac{1}{n-1} Y'_{(i)} (I_n - \frac{1}{n} E_n) Y_{(j)} \\ &= \frac{1}{n-1} \sum_{k=1}^n (y_{ki} - \frac{1}{n} y_{\cdot i}) (y_{kj} - \frac{1}{n} y_{\cdot j}). \end{aligned}$$

$\hat{\Sigma}$ hängt nicht von μ ab und es ist $\hat{\Sigma} = \hat{\Sigma}' \geq 0$.

Missing-Data-Problem

In Spalte i ($i=1, \dots, p$) werde nur $P_i Y_{(i)}$ beobachtet, wobei $P_i = \text{diag}(\Delta_{1i}, \dots, \Delta_{ni})$ mit $\Delta_{ki} \in \{0, 1\}$ (siehe Abschnitt 4.2. bzw. 1.2.0.)

Die orthogonale Projektion W_i auf $L_i := P_i L_0$ hat die Gestalt

$$W_i = \frac{1}{r_i} P_i E_n P_i \quad \text{mit } r_i := \text{Rg } P_i \quad (i=1, \dots, p),$$

$$\text{d.h. } W_i P_i Y_{(i)} = \frac{1}{r_i} Y_{0i} P_i (1, \dots, 1)' \quad (Y_{0i} := \sum_{k=1}^n \Delta_{ki} Y_{ki}),$$

$$\text{so da\ss } \tilde{\mu}_i := \frac{1}{r_i} Y_{0i}$$

$\forall \Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ lokal bester linearer erwartungstreuer Schätzer für μ_i ist ($i=1, \dots, p$).

$\tilde{\mu} := (\tilde{\mu}_1, \dots, \tilde{\mu}_p)'$ ist $\forall \Sigma = \Sigma' \geq 0$ gleichmäßig bester Schätzer für μ genau dann wenn die Bedingung (Δ) aus Abschnitt 4.2. erfüllt ist, d.h.

$$(\Delta) \quad \frac{1}{r_i} P_i P_j E_n P_j = \frac{1}{r_j} P_j P_i E_n P_i \quad \forall i, j=1, \dots, p.$$

Wegen $P_i P_j = P_j P_i$ und $P_i = \text{diag}(\Delta_{1i}, \dots, \Delta_{ni})$ folgt aus (Δ) insbesondere

$$\frac{1}{r_i} P_i P_j E_n P_i P_j = \frac{1}{r_j} P_i P_j E_n P_i P_j \quad \forall i, j=1, \dots, p$$

und damit $r_i = r_j \quad \forall i, j=1, \dots, p$ sowie $P_i P_j E_n P_j = P_j P_i E_n P_i$
 $\forall i, j=1, \dots, p.$

"Ersetzen fehlender Daten"

Offensichtlich ist $\tilde{\mu}_i = \frac{1}{n} (y_{\odot i} + \frac{n-r_i}{r_i} y_{\odot i}) = \frac{1}{n} z_{\cdot i}$

mit $z_{ki} := \Delta_{ki} y_{ki} + (1-\Delta_{ki}) \frac{n-r_i}{r_i} y_{\odot i} \quad (k=1, \dots, n; i=1, \dots, p).$

$x_{(i)} := \frac{1}{r_i} y_{\odot i} (I - P_i) (1, \dots, 1)'$ ist Lösung von

$$x = (I - P_i) \frac{1}{n} E_n (P_i y_{(i)} + x),$$

bzw. äquivalent:

$\frac{1}{r_i} y_{\odot i}$ ist Lösung von $x = \frac{1}{n} (y_{\odot i} + (n-r_i)x).$

Schätzung von $\Sigma = (\sigma_{ij})$

Die (bei $\mu_1 = \dots = \mu_p = 0$, $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ lokal) besten quadratischen erwartungstreuen Schätzer für σ_{ij} sind gegeben durch:

$$\begin{aligned} \tilde{\sigma}_{ij} &= \frac{(P_i Y(i))' (P_i P_j - W_i W_j) (P_j Y(j))}{\text{tr}(P_i P_j - W_i W_j) (P_j P_i - W_j W_i)} \\ &= \frac{\sum_{k=1}^n \Delta_{ki} \Delta_{kj} y_{ki} y_{kj} - \frac{1}{r_i r_j} y_{\odot i} y_{\odot j}}{r_{ij} \left\{ 1 - \frac{r_{ij}}{r_i \cdot r_j} \right\}} \quad (i, j = 1, \dots, p). \end{aligned}$$

$$(r_{ij} := \text{Rg}(P_i P_j) \quad i, j = 1, \dots, p).$$

Die (bei $\Sigma = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ lokal) besten invarianten erwartungstreuen quadratischen Schätzer für σ_{ij} sind gegeben durch:

$$\begin{aligned} \sigma_{ij}^R &= \frac{(P_i Y(i))' (P_i - W_i) (P_j - W_j) (P_j Y(j))}{\text{tr}(P_i - W_i) (P_j - W_j)} \\ &= \frac{\sum_{k=1}^n \Delta_{ki} \Delta_{kj} (y_{ki} - \frac{1}{r_i} y_{\odot i}) (y_{kj} - \frac{1}{r_j} y_{\odot j})}{r_{ij} \left\{ 1 - \frac{1}{r_i} - \frac{1}{r_j} + \frac{r_{ij}}{r_i \cdot r_j} \right\}} \quad (i, j = 1, \dots, p). \end{aligned}$$

$$\tilde{\sigma}_{ij} = \frac{\sum_{k=1}^n (z_{ki} - z_{\cdot i})(z_{kj} - z_{\cdot j})}{r_{ij} \left\{ 1 - \frac{1}{r_i} - \frac{1}{r_j} + \frac{1}{r_i r_j} \right\}}$$

mit z_{ki} wie oben ($k=1, \dots, n$; $i=1, \dots, p$).

Insbesondere ist also

$$\tilde{\sigma}_{ij}(Py) = \frac{(n-1) r_i r_j}{r_{ij}(r_i r_j - r_i - r_j + r_{ij})} \hat{\sigma}_{ij}(Py + x)$$

mit $x := (x'_{(1)}, \dots, x'_{(p)})'$, $x_{(i)}$ wie oben.

Bemerkung:

1) Für $i=j$ ($i=1, \dots, p$) ist

$$\begin{aligned} \tilde{\sigma}_{ii} &= \frac{1}{r_i - 1} \left(\sum \Delta_{ki} y_{ki}^2 - \frac{1}{r_i} y_{\ominus i}^2 \right) \\ &= \frac{1}{r_i - 1} \sum \Delta_{ki} (y_{ki} - y_{\ominus i})^2 = \tilde{\sigma}_{ii} . \end{aligned}$$

Wegen Theorem 3(3) gilt:

$\tilde{\sigma}_{ij} = \tilde{\sigma}_{ij} \quad \forall i, j=1, \dots, p \quad \forall y \in V$ genau dann wenn

$$\forall i, j=1, \dots, p : \left(P_i P_j E_n P_j = P_j P_i E_n P_i \text{ und } r_i = r_j \right)$$

- 2) Zur Berechnung von $\tilde{\sigma}_{ij}$ werden nur die Beobachtungspaare (y_{ki}, y_{kj}) ($k=1, \dots, n$) herangezogen, für die beide Meßwerte vorhanden sind, d.h. für die $\Delta_{ki} \Delta_{kj} \neq 0$ ist.

Im Gegensatz zu $\tilde{\sigma}_{ij}$ werden also Informationen vernachlässigt.

$\tilde{\sigma}_{ij}$ und $\text{Var}_P(\tilde{\sigma}_{ij})$ hängen explizit von μ_i und μ_j ab.

Falls jedoch $\mu_i = \mu_j = 0$ und $\sigma_{ij} = \delta_{ij} \sigma_{ii}$, so ist

$$\text{Var}_P(\tilde{\sigma}_{ij}) \leq \text{Var}_P(\tilde{\sigma}_{ij}).$$

- 3) Für $\Sigma^* := \frac{1}{n-1} Z' (I - \frac{1}{n} E_n) Z =: (\sigma_{ij}^*)$

$$\text{mit } Z := (p_1 y(1) + x(1), \dots, p_p y(p) + x(p))$$

gilt:

$$\tilde{\sigma}_{ij} = \frac{(n-1) r_i r_j}{r_{ij} (r_i r_j - r_i - r_j + r_{ij})} \sigma_{ij}^* \quad (i, j=1, \dots, p).$$

Σ^* ist der Wilks-Schätzer für Σ

(siehe WILKS 1932, TIMM 1970, GLEASON & STAELIN 1975):

Fehlende Daten werden durch Spaltenmittelwerte ersetzt.

Anmerkungen zu 4.1.- 4.3.

- 1) Eine notwendige und hinreichende Bedingung für die Existenz eines gleichmäßig besten linearen Schätzers für E_{pY} unter der Annahme des multivariaten Regressionsmodells $M(L, \theta)$ mit $L := (I_p \otimes X)$ und $\theta := \{I_n \otimes \Sigma \mid \Sigma = \Sigma' \geq 0\}$ - also $Q = I_n$ - sowie die Form des optimalen Schätzers $\hat{G} = \text{diag}(W_1, \dots, W_p)$ wurden bereits von Drygas angegeben.
(DRYGAS 1976, Theorem 3.3. mit Bemerkung 3.4. sowie Bemerkung 3.5.)

Einen Hinweis darauf, daß sich das multivariate Modell in p univariate Modell zerlegen läßt, findet sich ferner bei McDonald (MCDONALD, Theorem 1), dort im Zusammenhang mit dem Ersetzen fehlender Daten.

Die Existenz eines gleichmäßig besten quadratischen Schätzers für Σ im Fall einer vollständigen Datenmatrix wurde ebenfalls von Drygas nachgewiesen.

(DRYGAS 1977, Beispiel B)

- 2) Für den Fall unvollständiger Daten sind - insbesondere im Umfeld sozialwissenschaftlicher Fragestellungen - verschiedene Verfahren entwickelt worden, fehlende Werte zu ersetzen.

Ein allgemeiner Überblick über die Möglichkeiten des Umgangs mit Missing-Data in multivariaten Modellen findet sich bei MEYERS, Seite 331 ff.

Speziell im Hinblick auf das Schätzen der Kovarianzmatrix bzw. der Korrelationsmatrix gibt Timm verschiedene Methoden an, fehlende Daten zu ersetzen, u.a. durch Variablenmittelwerte. (TIMM 1970)

Dieses zuerst von Wilks empfohlene Vorgehen (WILKS 1932) findet meines Wissens breite Anwendung (SPSS enthält hierfür eine Option) - u.a. zur Berechnung von Startschätzern

in verbesserten mehrstufigen Verfahren (Hauptkomponentenverfahren und Regressionsverfahren - siehe GLEASON & STAELIN 1975, Seite 236 ff).

Für den Fall kleiner mittlerer Korrelationen sind jedoch mit diesen Verfahren gegenüber der Wilks-Methode keine Verbesserungen mehr zu erzielen.

(GLEASON & STAELIN 1976, Seite 250)

Es bleibt zu bemerken, daß man die korrekten (bei diagonalem Σ) optimalen invarianten Schätzer $\hat{\sigma}_{ij}$ aus den Wilks-Schätzern c_{ij}^* durch Multiplikation mit einem Korrekturfaktor erhält:

$$\hat{\sigma}_{ij} = c_{ij} \cdot \sigma_{ij}^* \quad (i, j=1, \dots, p)$$

wobei für das Modell $M(L, \theta)$ mit $L := (I_p \otimes X)$ und $\Theta := \{I_n \otimes \Sigma \mid \Sigma = \Sigma' \geq 0\}$

$$c_{ij} = \text{tr}(I - W_0) / \text{tr}(P_i - W_i)(P_j - W_j) \geq 1$$

für $i, j=1, \dots, p$ ist.

(vergleiche Absatz 1) der Anmerkungen zu 3.3., sowie MEYERS, Seite 332)

Speziell für $X := (1, \dots, 1)'$ ergibt sich

$$c_{ij} = (n-1) \left/ \frac{r_{ij}}{r_i r_j} \right. (r_i r_j - r_i - r_j + r_{ij})$$

(siehe Abschnitt 4.3., Bemerkung 3).

Schlußbemerkung

Die geometrische Betrachtungsweise des Gauß-Markov-Problems ermöglicht einen tieferen Einblick in die Bildungsgesetze linearer Schätzer. Insbesondere wird ein neuer Zugang zum Missing-Data-Problem ("Ersetzen fehlender Daten") eröffnet, der zu einer vollständigen Klärung der Situation bei beliebiger Kovarianz führt. Von besonderer Bedeutung ist die Begründung einer einheitlichen Theorie für lineare und quadratische Schätzer.

Da in den vorliegenden Ausführungen die endliche Dimensionalität des Beobachtungsraumes im Wesentlichen unberücksichtigt bleibt, bietet sich eine Verallgemeinerung der Ergebnisse auf Hilbertraum-wertige Zufallsvariable an (DRYGAS 1975, REHDER 1980).

Im Hinblick auf mögliche Anwendungen könnte das hier angegebene Verfahren zur Entwicklung quadratischer Schätzfunktionen für den Fall unvollständiger Daten von Interesse sein sowie der Hinweis auf die Berechnungsmöglichkeit bester invarianter Schätzer mit Hilfe eines aufgefüllten Datensatzes.

Insbesondere wäre zu prüfen, wie sich die Verwendung der angegebenen Korrekturfaktoren für die empirischen Kovarianzen einer multivariaten Stichprobe in verschiedenen statistischen Analyseverfahren (z.B. Faktoren-Analyse) auswirkt.

Zeichenerklärung

$(\cdot, \cdot), (\cdot, \cdot)_0$ Skalarprodukte (auf V)

Für eine Teilmenge $M \subseteq V$:

$[M]$ lineare Hülle von M in V

Für lineare Teilräume $M, N \subseteq V, M_i \subseteq V (i \in I)$:

$M + N, \sum_{i \in I} M_i$ Summe

$M \oplus N$ direkte Summe

$M \dot{\oplus} N$ orthogonale Summe

$M^\perp, M^{\perp_0}$ orthogonales Komplement von M in V bezüglich $(\cdot, \cdot)$ bzw. $(\cdot, \cdot)_0$.

P_M, P_M^0 orthogonale Projektion auf M bezüglich $(\cdot, \cdot)$ bzw. $(\cdot, \cdot)_0$.

$\mathcal{L}(V, V)$ Menge der linearen Abbildungen von V in sich

$S(V)$ Menge der symmetrischen linearen Abbildungen von V in sich

$S_+(V)$ Menge der symmetrischen, positiv semi-definiten linearen Abbildungen von V in sich

$V \otimes V$ Tensorprodukt

Für lineare Abbildungen A, B :

A' Adjungierte zu A

AM Bild von M unter A

$K(A)$ Kern von A , Nullraum

$Rg(A)$ Rang von A

$\text{tr } A$ Spur von A

A^- g-Inverse (nicht eindeutig)

A^+ Moore-Penrose-Inverse (eindeutig)

$R_{n,n}$ Menge der reellen $n \times n$ -Matrizen
 I_n n -te Einheitsmatrix
 E_n $(a_{ij})_{i,j=1,\dots,n}$ mit $a_{ij}=1 \quad \forall i,j$
 O_n Nullmatrix

Für $A, B \in R_{n,n}$:

$A \otimes B$ Kronecker-Produkt

Literaturverzeichnis

- (1) Allan, F.E. & Wishart, J. (1930):
A Method of Estimating the Yield of a Missing
Plot in Field Experimental Work
J.Agric.Sci. 20, 399-406
- (2) Albert, A. (1973):
The Gauss-Markov-Theorem for Regression
Models with Possibly Singular Covariances
Siam J.Appl.Math. 24, 182-187
- (3) Anderson, R.L. (1946):
Missing Plot Techniques
Biometrics 2, 41-47
- (4) Biggers, J.D. (1959):
The Estimation of Missing and Mixed-up
Observations in Several Experimental Designs
Biometrika 46, 91-105
- (5) De Lury, D.B. (1946):
The Analysis of Latin Squares when some
Observations are Missing
J.Am.Statist.Ass. 41, 370-389
- (6) Draper, N.R. & Stoneman, D.M. (1964):
Estimating Missing Values in Unreplicated
Two-Level Factorial and Fractional Factorial
Desings
Biometrics 20, 443-458

- (7) Drygas, H. (1969):

On a Generalization of the Farkas' Theorem
Unternehmensforschung 13, 283-290

- (8) Drygas, H. (1970):

The Coordinate-Free Approach to Gauss-Markov-
Estimation
Lecture Notes in Operations Research and
Mathematical Systems, Springer

- (9) Drygas, H. (1971):

Consistency of the Least Squares and Gauss-
Markov Estimators in Regression Models
Z. Wahrscheinlichkeitstheorie verw. Geb. 17,
309-326

- (10) Drygas, H. (1972):

The Estimation of Residual Variances in
Regression Analysis
Math. Operationsforsch. Statist. 3, 373-388

- (11) Drygas, H. (1975):

Estimation and Prediction for Linear Models
in General Spaces
Math. Operationsforsch. Statist. 6, 301-324

- (12) Drygas, H. (1976):

Gauss-Markov Estimators for Multivariate
Linear Models with Missing Observations
Ann. Statist. 4, 779-787

- (13) Drygas, H. (1977):

Best Quadratic Unbiased Estimation in

Variance-Covariance Component Models

Math. Operationsforsch. Statist., Ser. Statist. 8,
211-231

- (14) Durbin, J. & Kendall, M.G. (1951):

The Geometry of Estimation

Biometrika 38, 150-158

- (15) Eaton, M.L. (1970):

Gauss-Markov Estimation for Multivariate

Linear Models - A Coordinate-Free Approach

Ann. Math. Statist. 41, 528-538

- (16) Eaton, M.L. (1978):

A Note on the Gauss-Markov Theorem

Ann. Inst. Statist. Math. 30, 181-184

- (17) Gleason, T.C. & Staelin, R.A. (1975):

A Proposal for Handling Missing Data

Psychometrika 40, 229-252

- (18) Halmos, P. (1958):

Finite Dimensional Vector Spaces

Springer 1958, 1974

- (19) Kleffe, J. (1975):

Invariant Methods for Estimating Variance

Components in Mixed Linear Models

Math. Operationsforsch. Statist., Ser. Statist. 8,
233-250

- (20) Kleffe, J. (1976):

A Note on MINQUE for Normal Models
Math. Operationsforsch. Statist. 7, 707-714

- (21) Kleffe, J. (1980):

On Recent Progress of MINQUE-Theory -
Nonnegative Estimators, Consistency, Asymptotic
Normality and Explicite Formulae
Math. Operationsforsch. Statist. 11, 563-588

- (22) Kleffe, J. & Pincus, R. (1974):

Bayes and Best Quadratic Unbiased Estimators
for Parameters of the Covariance Matrix in
a Normal Linear Model
Math. Operationsforsch. Statist. 5, 43-67

- (23) Kruskal, W. (1961):

The Coordinate-Free Approach to Gauss-Markov
Estimation and its Application to Missing
and Extra Observations
Fourth Berkeley Symp. Math. Statist. Prob. 1,
435-451

- (24) Kruskal, W. (1968):

When are Gauss-Markov and Least-Squares
Estimators Identical?
A Coordinate-Free Approach
Ann. Math. Statist. 39, 70-75

- (25) Kruskal, W. (1975):

The Geometry of Generalized Inverses
J. Roy. Statist. Soc., Ser. B, 37, 272-283

- (26) LaMotte, L.R. (1973):

Quadratic Estimation of Variance Components
Biometrics 29, 311-330

- (27) McDonald, L. (1971):

On the Estimation of Missing Data in the
Multivariate Linear Model
Biometrics 27, 535-543

- (28) Meyers, E.D. (1973):

Time-Sharing Computation in the Social
Sciences
Prentice-Hall

- (29) Mitra, S.K. & Moore, B.J. (1973):

Gauss-Markov Estimation with an Incorrect
Dispersion Matrix
Sankhya, Ser.A 35, 139-152

- (30) Nie, E.D. et.al.

Statistical Package for the Social Sciences
(SPSS)
McGraw Hill

- (31) Puckelsheim, F. (1976):

Estimating Variance Components in Linear Models
J.Mult.Anal. 6, 626-629

- (32) Puckelsheim, F. (1977):

On HSU's Model in the Regression Analysis
Math. Operationsforsch. Statist., Ser. Statist. 8,
323-331

- (33) Rao, C.R. (1971) a):

Unified Theory of Linear Estimation
Sankhya, Ser. A 33, 371-393

- (34) Rao, C.R. (1971) b):

Minimum Variance Quadratic Unbiased Estimation
of Variance Components
J. Mult. Anal. 1, 445-456

- (35) Rao, C.R. (1972):

Unified Theory of Least Squares
Commun. in Statist. 1, 1-8

- (36) Rao, C.R. (1973):

Representations of Best Linear Unbiased
Estimation in the Gauss-Markov Model with
a Singular Dispersion Matrix
J. Mult. Anal. 3, 276-292

- (37) Rehder, W. (1980):

When is the Pseudo-Best Estimator BLUE?
Ann. Inst. Statist. Math. 32, 179-185

- (38) Schönfeld, P. (1967):

Generalized Best Linear Unbiased Estimation
CORE Discussion Paper No. 6713, Louvain

- (39) Seely, J. (1971):

Quadratic Subspaces and Completeness

Ann.Math.Statist. 42, 710-721

- (40) Swallow, W.H. & Searle, S.R. (1978):

Minimum Variance Quadratic Unbiased Estimation
of Variance Components

Technometrics 20, 265-272

- (41) Thompson, H.R. (1956):

Extensions to Missing Plot Techniques

Biometrics 12, 241-244

- (42) Timm, N.H. (1970):

The Estimation of Variance Components
and Correlation Matrices from Incomplete Data

Psychometrika 35, 417-438

- (43) Townsend, E.C. & Searle, S.R. (1971):

Best quadratic Unbiased Estimation of
Variance Components from Unbalanced Data in
the One-Way Classification

Biometrics 27, 643-658

- (44) Wilkinson, W.L. (1958) a):

Estimation of Missing Values for the Analysis
of Incomplete Data

Biometrics 14, 257-286

(45) Wilkinson, W.L. (1958) b):

The Analysis of Variance and Deviation of
Standard Errors for Incomplete Data
Biometrics 14, 360-384

(46) Yates, F. (1933):

The Analysis of Replicated Experiments when
the Field Results are Incomplete
Emp. J. Exp. Agric. 1, 129-142

Lebenslauf

17.10.1950	als Tochter des Landgerichtsrates Dr. Heinrich Buchholz und seiner Frau Ursula, geb. Worringer, in Essen geboren
1957 - 1961	Besuch der katholischen Volksschule in Essen
1961 - 1969	Besuch des neusprachlichen Gymnasiums in Essen
Juni 1969	Abitur
1969 - 1971	Studium der Mathematik und Philosophie in Göttingen
seit 1970	verheiratet mit dem Arzt Richard Malin
Mai 1971	Geburt unserer Tochter Ruth
1971 - 1976	Studium der Mathematik in Hannover (Nebenfach: Physik)
November 72	Diplom-Vorprüfung im Fach Mathematik
1974 -1975	Studentische Hilfskraft
August 1976	Diplom-Hauptprüfung im Fach Mathematik (Schwerpunktfach: Stochastik)
1976 - 1979	Wissenschaftliche Hilfskraft an der Universität Essen
1979 - 1982	Verwalter der Stelle eines wissenschaft- lichen Assistenten in Essen
Juli 1982	Geburt unserer Tochter Lydia

